Power-Proportional Optical Links

Abdullah Ibn Abbas

A Thesis

In the Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Electrical and Computer Engineering) at

Concordia University

Montreal, Quebec, Canada

March 2021

© Abdullah Ibn Abbas, 2021

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Abdullah Ibn Abbas

Entitled: Power-Proportional Optical Links

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Electrical & Computer Engineering)

complies with the regulation of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

		Chair
Dr. Roch Gl	itho	
		External Examiner
Dr. Masum	Hossain	
		External to Program
Dr. Rajagop	alan Jayakumar	
		Examiner
Dr. Asim J.	Al-Khalili	
		Examiner
Dr. Chunyar	n Wang	
		Thesis Supervisor
Dr. Glenn C	owan	
Approved by		
11 J	Dr. Wei-Ping Zhu, Graduate Pro	ogram Director
March 09, 2021		
	Dr. Mourad Debbabi, Dean	
	Gina Cody School of Engineerin	g and Computer Science

ABSTRACT

Power-Proportional Optical Links

Abdullah Ibn Abbas, Ph.D. Concordia University, 2021

The continuous increase in data transfer rate in short-reach links, such as chip-to-chip and between servers within a data-center, demands high-speed links. As power efficiency becomes ever more important in these links, power-efficient optical links need to be designed.

Power efficiency in a link can be achieved by enabling power-proportional communication over the serial link. In power-proportional links, the power dissipated by a link is proportional to the amount of data communicated. Normally, data-rate demand is not constant, and the peak data-rate is not required all the time. If a link is not adapted according to the data-rate demand, there will be a fixed power dissipation, and the power efficiency of the link will degrade during the sub-maximal link utilization. Adapting links to real-time data-rate requirements reduces power dissipation. Power proportionality is achieved by scaling the power of the serial link linearly with the link utilization, and techniques such as *variable data-rate* and *burst-mode* can be adopted for this purpose. Links whose data rate (and hence power dissipation) can be varied in response to system demands are proposed in this work.

Past works have presented rapidly reconfigurable bandwidth in *variable data-rate* receivers, allowing lower power dissipation for lower data-rate operation. However, maintaining synchronization during reconfiguration was not possible since previous approaches have introduced changes in front-end delay when they are reconfigured. This work presents a technique that allows

rapid bandwidth adjustment while maintaining a near-constant delay through the receiver suitable for a power-scalable variable data-rate optical link. Measurements of a fabricated integrated circuit (IC) show nearly constant energy per bit across a 2× variation in data rate while introducing less than 10 % of a unit interval (UI) of delay variation.

With continuously increasing data communication in data-centers, parallel optical links with ever-increasing per-lane data rates are being used to meet overall throughput demands. Simultaneously, power efficiency is becoming increasingly important for these links since they do not transmit useful data all the time. The burst-mode solution for vertical-cavity surface-emitting laser (VCSEL)-based point-to-point communication can be used to improve links' energy efficiency during low link activity. The burst-mode technique for VCSEL-based links has not yet been deployed commercially. Past works have presented burst-mode solutions for single-channel receivers, allowing lower power dissipation during low link activity and solutions for fast activation of the receivers. However, this work presents a novel technique that allows rapid activation of a front-end and fast locking of a clock-and-data-recovery (CDR) for a multi-channel parallel link, utilizing opportunities arising from the parallel nature of many VCSEL-based links. The idea has been demonstrated through electrical and optical measurements of a fabricated IC at 10 Gbps, which show fast data detection and activation of the circuitry within 49 UIs while allowing the front-end to achieve better energy efficiency during low link activity. Simulation results are also presented in support of the proposed technique which allows the CDR to lock within 26 UIs from when it is powered on.

ACKNOWLEDGMENTS

This thesis is the outcome of the prayers and support from my beloved parents and teachers who encouraged me to carry out this work.

I would like to express my sincere gratitude to Dr. Glenn Cowan for his supervision, constructive suggestions, and guidance throughout the course of this work. I am also highly indebted to him for providing financial support throughout the degree, without which it would have been challenging to carry out the research smoothly. Also, his in-depth knowledge in the field paved the path for me to learn and add to my knowledge many practical aspects of the encountered issues.

I would further extend my thanks to Dr. Asim J. Al-Khalili, Dr. Chunyan Wang, Dr. Rajagopalan Jayakumar, and Dr. Masum Hossain for reading this thesis and providing useful comments and suggestions to improve it.

It is worth recognizing and thanking our group members Chris Williams (Ph.D. graduate) and Diaaeldin Abdelrahman for helpful discussion during this work. Special thanks to Xiangdong Jia for helping me with the layout and simulations during my first tapeout.

I am extremely grateful to Professor Odile Liboiron-Ladouceur of McGill for allowing me to use the optical testbench to perform optical measurements, Dr. Reza for helping me establish the optical test setup, and Professor Steve Shih at Concordia University for generously sharing his probe station. When mentioning the optical measurements, it is unimaginable not to mention Dr. Rolston of Smiths Interconnect (Reflex Photonics) for his generous offer of wirebonding and supply of the photodiodes. I am also grateful to Ted Obuchowicz for his prompt technical and other non-technical support during my degree. Further, an unforgettable contribution from Canadian Microelectronic Corporation (CMC) relating to fabrication, CAD tools, test equipment, etc. A deep appreciation and thanks to James Millar and James Dietrich of CMC for their promptness in managing the test equipment and invaluable support without which the measurement would have been incomplete.

I am further grateful for all the valuable help received from the administrative staff of the Department of Electrical and Computer Engineering during my stay.

Finally, I would not forget to thank all my relatives and friends whose sincere prayers and kind words are always the source of my energy. Concluding with thanks to all who contributed directly and indirectly to this piece of work.

TABLE OF CONTENTS

List of Figures	xi
List of Tables	xxi
List of Acronyms	xxii
Chapter 1 – Introduction	1
1.1 Motivation	1
1.2 Objectives/Problems	3
1.3 Claim of Originality	4
1.4 Publications and Contributions	5
1.5 Thesis Organization	7
Chapter 2 – Background and Literature Review	8
2.1 Wireline Communication System	8
2.2 Transimpedance Amplifier (TIA)	10
2.3 Post-Amplifier	11
2.4 VCSEL	12
2.5 Photodiode	13
2.6 Bit Error Rate (BER) and Q-factor	14
2.7 Extinction Ratio and Optical Modulation Amplitude (OMA)	
2.8 Sensitivity	23
2.9 Eye Diagram	25
2.10 Transfer Function and Group Delay	
2.11 Phase Noise and Jitter	
2.12 Energy Proportional Links	
2.13 Gear-Shifting Link/Variable-Rate Link	

2.14 Burst-Mode Link	
2.15 Phase Interpolator-based CDR	
2.16 Literature Review	
2.16.1 FE and CDR for Gear-Shifting Link	
2.16.2 FE and CDR for Burst-Mode Link	45
2.17 Summary	
Chapter 3 – Rapidly Reconfigurable Variable-Rate Optical Link Receiver	53
3.1 Receiver Design	53
3.2 Power-Scalable, Variable-Rate and Constant-Delay Front-End Architecture	55
3.3 Front-End Design Methodology	59
3.4 CDR Architecture	67
3.5 Simulation and Measurement Results	75
3.5.1 S-parameters and Gain	76
3.5.2 Matched-Delay through the FE	79
3.5.3 Bit Errors during Mode Transition	
3.5.4 FE Power Dissipation vs Bandwidth	
3.5.5 Sensitivity vs Data-rate	
3.5.6 Clock Performance	
3.6 Comparison of the Variable-Rate FE with Published Works	
3.7 Conclusion	89
Chapter 4 – Burst-Mode CDR for VCSEL-Based Parallel Optical Links	91
4.1 Phase Relationship of Data over Parallel Channels	92
4.2 Proposed Receiver Architecture	95
4.2.1 Proposed Burst-Mode Link Protocol	
4.2.2 Sense Circuit Description	

4.2.3 CDR Architecture and Description	101
4.2.4 CDR Loop Dynamics, Stability and Inter-Lane Tracking	106
4.3 Simulation Results	110
4.4 Comparison of the Burst-Mode CDR with Published Works	120
4.5 Conclusion	122
Chapter 5 – Burst-Mode Front-End for VCSEL-Based Parallel Optical Links	123
5.1 Proposed Receiver Architecture	124
5.1.1 Proposed Burst-Mode Link Protocol	125
5.1.2 Front-End Architecture and Description	128
5.2 FE Measurements with Electrical Input Signal	131
5.2.1 S-parameters and Gain	132
5.2.2 Analog Input Voltage Sensitivity and Bathtub Curve	133
5.2.3 Turn-on Time for Burst-mode Operation	135
5.3 FE Measurements with Wire-Bonded Photodiode	136
5.4 Comparison of the Burst-Mode FE with Published Works	138
5.5 Conclusion	139
Chapter 6 – Conclusions	141
6.1 Thesis Highlights	141
6.2 Future Work	143
6.2.1 Variable-Rate Receiver	143
6.2.2 Burst-Mode Multi-Channel Parallel Receiver Architecture and Related Issues	s 143
6.2.3 Burst-Mode ILO	144
6.2.4 Re-Designing of the Variable-Rate and the Burst-Mode CDR	149
Appendix	150

A Injection Locked Oscillator (ILO) Design for Variable-Rate CDR	
B Injection Locked Oscillator (ILO) Design for Burst-Mode CDR	
C On-chip Shift Register Design for Control Bits	
References	

List of Figures

Figure 1.1	Power breakdown of a typical short reach optical link	2
Figure 2.1	Typical communication links: (a) electrical and (b) optical	8
Figure 2.2	Conventional optical front-end architecture	9
Figure 2.3	(a) Inverter-based shunt feedback TIA. (b) TIA with photodiode model at the input and load at the output	10
Figure 2.4	Cherry-Hooper inverter-based post-amplifier	12
Figure 2.5	(a) Symbolic representation of a laser. (b) Micrograph of a commercial (850-nm) VCSEL (an array of four lasers) (Finisar)	13
Figure 2.6	Symbolic representation of a photodiode. (b) Micrograph of a commercial (850-nm) photodiode (Cosemi)	14
Figure 2.7	Probability of error for binary signal	16
Figure 2.8	Probability of error for a binary signal with greater separation between the 'high' and 'low' values	16
Figure 2.9	BER vs Q-factor	20
Figure 2.10	Non-return to zero (NRZ) data showing the power level	22
Figure 2.11	(a) Time domain random NRZ data pattern. (b) Slices of the data in (a) at an interval of bit period to create an eye diagram	26
Figure 2.12	Eye diagram of the data shown in Figure 2.10. (a) Eye diagram corresponding to the slicing interval over one bit period or one unit interval (UI) by selecting	

	the start time to position the eye in the center. (b) Eye diagram corresponding	
	to the slicing interval over two bit periods and positioning it in the center with	
	a proper start time. (c) Eye diagram as presented in (a) but with a different start	
	time which leads to a circular shift in the eye. (d) Eye diagram as presented in	
	(b) but with a different start time which leads to a circular shift in the	77
	eye	21
Figure 2.13	A practical eye diagram with slicing interval over two bit periods and the one	
	eye positioned in the center	28
Figure 2.14	Activity of a gear-shifting link as a function of time	30
Figure 2.15	Ideal power dissipation and energy-per-bit of a gear-shifting link [7]	30
Figure 2.16	Conceptual diagram of (a) a conventional link, and (b) a power-proportional	
	burst-mode link, illustrating the power dissipation as a function of link	
	utilization	31
Figure 2.17	Standard phase interpolator block diagram	32
Figure 2.18	Variable-rate and power scalable front-end [16]	36
Figure 2.19	Energy efficiency plot as a function of data rate for the case presented in	
	[16]	37
Figure 2.20	Response time of the front-end [16]	38
Figure 2.21	Rate selection mechanism. Solid circles represent sample points and hollow	
	circles represent discarded samples [15]	39
Figure 2.22	Block diagram of PLL with VCO-bank used in [14]	40
Figure 2.23	Open loop and closed loop phase noise plot comparing the number of	
	connected VCOs [14]	41
Figure 2.24	Phase excursion plot for dynamic activation of sub-VCO [14]	41

Figure 2.25	Data streams showing the input data at front-end, and (a) the data at the output	
	of a conventional front-end for different bandwidth settings where $x \neq y$ (b)	
	the data at the output of a fixed delay front-end for different bandwidth settings	43
Figure 2.26	Data eyes showing the input data of the front-end, and (a) the data at the output	
	of a conventional front-end for different bandwidth settings where $x \neq y$ (b) the	
	data at the output of a fixed delay front-end for different bandwidth settings	43
Figure 2.27	A conceptual representation of a passive optical network (PON)	46
Figure 2.28	(a) Graphical representation of transients in a burst-mode link and (b) conceptual illustration of response times as a function of off-state power	47
Figure 2.29	Proposed pulse to reduce VCSEL wake-up time [19]. (a) Start preamble with some ns of logic'1's, which will double VCSEL current compared with preamble 101010. (b) A short high-amplitude current pulse is supplied as a wake-up pulse prior to the preamble	48
Figure 2.30	(a) Analog front-end schematic, (b) burst-mode CDR block diagram, and (c) phase-locking dynamics as presented in [23]	49
Figure 2.31	(a) Link protocol and a burst-mode optical receiver. (b) Power-on time of the receiver [28]	50
Figure 2.32	(a) Burst-mode optical receiver. (b) Transient behavior (simulated) of the receiver [29]	51
Figure 3.1	Block diagram of the proposed receiver incorporating the architecture of the variable bandwidth and constant delay front-end	54

Figure 3.2	Variable-rate data stream and the sampling clock phases. Shaded clock phases are unavailable during low data rate	55
Figure 3.3	(a) Transient (dashed for input and solid for output data) and (b) group-delay plots for the delay through the front-end. Grey and black traces for 2.5 GHz and 5 GHz bandwidth respectively	58
Figure 3.4	Simulated bandwidth, gain, and the group delay of a front-end against the normalized value of the feedback resistor (R _a) with (a) TIA and three-stage post-amplifier, (b) TIA and a single-stage post-amplifier	61
Figure 3.5	Circuit level simulation results of bandwidth, gain, and group delay for an inverter-based front-end with a different number of post-amplifier stages against the scaling factor of the feedback resistors R_a and R_b	63
Figure 3.6	Skewing of R_a and R_b through scaling factors 'a' and 'b' to adjust the group delay in Figure 3.5 using one post-amplifier stage	65
Figure 3.7	Simulated group delay of the designed front-end for low and high bandwidth modes at 1.0 V supply voltage. Grey and black traces for 2.5 GHz and 5 GHz bandwidth, respectively	66
Figure 3.8	Dual-loop CDR	68
Figure 3.9	Architecture of the proposed CDR	69
Figure 3.10	Phase noise and power reconfigurable eight-stage ILO. Shaded blocks are OFF at low data rate mode	69
Figure 3.11	Proposed reconfigurable phase rotator architecture. Shaded blocks are powered down at low data-rate operation	71
Figure 3.12	Functional block diagram of the digital loop filter and the S-bit generation sequence for (a) high data rate (10 Gbps) and (b) low data rate (5 Gbps)	73

Figure 3.13	(a) Optical test setup for BER and eye-diagram measurements. (b) ENEPIG finished test board with wirebonded PD and die. (c) Die photo of the receiver chip in 65 nm CMOS	75
Figure 3.14	Measured at 1.0 V: (a) S-parameters, (b) transimpedance gain, and (c) group delay of the FE	78
Figure 3.15	(a) Simulated delay mismatch due to process variation against temperature (i) and in the presence of $\pm 5\%$ supply voltage variation at 27°C for TT process (ii). (b) Optically measured delay mismatch at supply voltages of 0.95 V (i) and 1.05 V (ii). (c) Optically measured overlaid eye-diagrams at low data-rates showing matched delay through the FE at two bandwidth settings for a supply voltage of 1.0 V.	80
Figure 3.16	Analog output data, output DC voltage, and the number of bit errors during reconfiguration. (a) Transition from low bandwidth to high bandwidth mode. (b) Transition from high bandwidth to low bandwidth mode	82
Figure 3.17	Power dissipation as a function of bandwidth for different supply voltages	83
Figure 3.18	Measured current sensitivity for different data-rates at low and high FE modes with optical input	84
Figure 3.19	Optically measured eye-diagrams at 4 Gbps (a) and 8 Gbps (b) at sensitivity level optical input powers	84
Figure 3.20	Measured oscilloscope plot of the injection locked clock. (a) low-mode operation and (b) high-mode operation	85
Figure 3.21	Phase noise plots obtained during high and low-mode operations	86
Figure 3.22	(a) Measured clock output during mode switching and (b) sudden phase drift of the clock with settling time to achieve final static-phase deskew	87

Figure 4.1	Phase relationship between the data of two channels in the presence of a ppm	
	frequency offset between data streams and the receiver's reference clock	93
Figure 4.2	(a) Phase relationship of two data streams, their recovered clocks and the receiver's reference clock, shown at the end of a burst and the start of the next burst on the burst-mode channel. Shown when the burst-mode channel's PR code is not updated. (b) The aforementioned phase relationship with the proposed proxy timing recovery scheme where the PR code of the burst-mode channel is updated between bursts.	95
Figure 4.3	(a) A conventional twelve-channel parallel optical link architecture.(b) Conceptual block diagram of the proposed burst-mode receiver for a twelve-channel parallel optical link showing one always-on channel and eleven burst-mode channels.	96
Figure 4.4	A proposed two-channel burst-mode receiver consisting of one always-on channel and one burst-mode channel	98
Figure 4.5	(a) Proposed link protocol for the burst-mode receiver. (b) Sampling clock and possible incoming data phases relative to the clock. (c) Sampled data indicating positive edge transitions	99
Figure 4.6	Circuit details of the SENSE CKT incorporating "Off Detect" and "Burst Detect"	101
Figure 4.7	A typical phase interpolator-based CDR	102
Figure 4.8	Block diagram of the CDR architecture of the proposed two-channel parallel optical link	103
Figure 4.9	Circuit details of the loop filter and the graphical illustration of the generation of UP/DN and TRIG signals	104

Figure 4.10	(a) PR circuit with transistor level diagram of one cell. (b) PR control logic generating 64 control bits for the PR. (c) Block diagram of the 8-phase generator incorporating three cascaded ILOs for the generation of eight clock	105
	phases with 45° phase spacing	105
Figure 4.11	Block diagram of the implemented CDR used for analysis	107
Figure 4.12	Layout of the CDR in 65 nm CMOS process	110
Figure 4.13	Theoretical and simulated JTOL of the proposed CDR	111
Figure 4.14	Phase delay plot of the phase rotator against phase steps	111
Figure 4.15	Mismatch variation at phase rotator code with (a) minimum phase delay error, (b) maximum phase delay error	112
Figure 4.16	Process variation at phase rotator code with (a) minimum phase delay error, (b) maximum phase delay error	112
Figure 4.17	(a) Phase delay variation as a function of run number with PR codes resulting in phase step number 0 and phase step number 16. (b) Phase delay difference between these two PR codes as a function of run number	113
Figure 4.18	Tolerated uncorrelated jitter by the proxy timing recovery plotted against the idle period with three different jitter periods of 1000 ns, 242 ns, and 121 ns	114
Figure 4.19	Simulated (extracted) locking behavior of the complete CDR	115
Figure 4.20	Analog output data and sampling clock of burst-mode and always-on channels with ON-OFF signal. Enlarged view of areas before and after the idle-state (40 ns) are also presented in this figure and subsequent figures. The enlarged view is for sampling behavior in the presence of 1000 ppm frequency offset between the transmit and receive clock frequency and without proxy timing recovery scheme. Burst-mode channel-2 also includes the nonlinear effect of the PR.	

	Without the use of proxy timing-recovery, the clocks of the burst-mode	
	channel are not locked at the end of the preamble bits	117
Figure 4.21	The simulation is repeated with proxy timing recovery scheme and only the	
	data and the sampling clocks of the burst-mode channels are shown because	
	the data and the sampling clock of the always-on channel always remain in	
	lock condition after certain period of initialization. The figure shows locking	
	within the preamble period of 49 UI in the presence of 1000 ppm frequency	
	offset (burst-mode channel-2 includes the PR nonlinearity)	118
Figure 4.22	The simulation is with proxy timing recovery in the presence of 300 ppm	
	frequency offset (burst-mode channel-2 includes the PR nonlinearity) and 4.14	
	MHz correlated jitter of amplitude 0.5 UI_{p-p} . The clocks of the burst-mode	
	channels are found to be locked successfully within the preamble period	118
Figure 4.23	Calculated phase difference (Δt) between the data and the sampling clock	
	resulting due to off period	119
Figure 5.1	A proposed two-channel burst-mode receiver consisting of one always-on	
	channel and one on-off channel	124
Figure 5.2	(a) Proposed link protocol for fast turn-on burst-mode FE. (b) Sampling clock	
	and possible incoming data phases relative to the clock. (c) Sampled data	
	indicating positive edge transitions	126
Figure 5.3	Block diagram of the proposed receiver incorporating the architecture of the	
	variable bandwidth front-end	127
Figure 5.4	Circuit details of the SENSE CKT incorporating "transition from 0 to 1	
	detection"	127
Figure 5.5	Simulated settling behavior of the FE (without the input data): (a) while	
	turning on and (b) while turning off	130

Figure 5.6	(a) Simulated transfer function with low frequency cut-off. (b) Simulated transient with three positive-edge detection turn-on time and DC-level.	
	circuit	130
Figure 5.7	(a) Electrical test setup for S-parameter, BER and measurements with probed input and output.(b) ENEPIG finished test board with wirebonded PD and die.(c) Die photo of the receiver chip in 65 nm CMOS	132
Figure 5.8	Measured at 0.95 V: (a) S-parameters and (b) transimpedance gain of the FE.	133
Figure 5.9	Measured (a) voltage sensitivity for different data rates at high and low bandwidths of the analog FE with electrical input and (b) bathtub curves at 1.5x sensitivity level electrical inputs	134
Figure 5.10	Measured eye-diagrams at (a) 10 Gbps with high bandwidth and (b) 2.0 Gbps low bandwidth at sensitivity level electrical inputs	134
Figure 5.11	Output transient behavior of the FE with burst of input data (Electrical input). (a) Multiple data burst, (b) Initialization bits followed by PRBS 2 ⁷ -1 pattern in a single burst	135
Figure 5.12	Optical test setup for BER and eye-diagram measurements with probed output.	136
Figure 5.13	Measured (a) current sensitivity for different data rates at high and low bandwidths of the analog FE with optical input and (b) bathtub curves at 1.5x sensitivity level optical inputs	137
Figure 5.14	Measured eye-diagrams at (a) 10 Gbps with high bandwidth and (b) 2.0 Gbps low bandwidth at sensitivity level optical inputs	137
Figure 6.1	Inverter-based five-stage injection locked ring oscillator	146

Figure 6.2	Simulated overlapped time periods of the clock signal for determining the lock	
	time of the inverter-based ring oscillator	147
Figure 6.3	NAND gate-based five-stage injection locked ring oscillator	148
Figure 6.4	Simulated overlapped time periods of the clock signal for determining the lock	
	time of the NAND gate-based ring oscillator	148
Figure A.1	Eight-stage ring oscillator using a cross-coupled, pseudo-differential current-	
	starved structure	150
Figure B.1	Four-stage ring oscillator using a cross-coupled, pseudo-differential current-	
	starved structure	152
Figure C.1	An N-bit shift register showing its input and output connections	154
Figure C.2	N-bit shift register consisting of small and identical sub-blocks	155
Figure C.3	Connection of DFF in each sub-block of the N-bit shift register	155

List of Tables

Table-2.1	Data rate-dependent wait-time for10-12 BER	21
Table-2.2	Overall receiver front-end performance summary [16]	37
Table-3.1	Group delays with different number of post-amplifier stages	64
Table-3.2	FE performance summary	79
Table-3.3	FE performance comparison	89
Table-4.1	Power breakdown of the CDR	110
Table-4.2	CDR performance comparison	121
Table-5.1	FE performance summary	133
Table-5.2	FE Performance comparison	139

List of Acronyms

additive white Gaussian noise
bit error ratio
bit error rate tester
clock and data recovery
complementary metal oxide semiconductor
delay-locked loop
device under test
error detector
electrostatic discharge
front-end
high bandwidth
high mode
input-output
injection-locked oscillator
inter-symbol interference
jitter tracking bandwidth
jitter tolerance
limiting amplifier
low bandwidth
loop filter
low mode
main amplifier
majority voting
non-return-to-zero
offset compensation
optical line terminal
optical modulation amplitude

ONU	optical network unit
PA	post-amplifier
PC	polarization controller
PD	photodiode
PI	phase interpolator
PLL	phase-locked loop
PON	passive optical network
PPG	pulse pattern generator
PR	phase rotator
PRBS	pseudorandom binary sequence
PVT	process, voltage, and temperature
QFN	quad flat no-leads
Rx	receiver
SMA	sub-miniature version a
SNR	signal-to-noise ratio
TDM	time division multiplexing
TIA	transimpedance amplifier
Tx	transmitter
UI	unit interval
VCO	voltage controlled oscillator
VCSEL	vertical cavity surface-emitting laser
VGA	variable-gain amplifier
VNA	vector network analyzer

In this chapter, trends in high-speed interconnects are highlighted, and the motivation behind the present work is discussed. Also, the existing problems that need to be addressed are presented along with the highlights of the work done in this dissertation.

1.1 Motivation

Present high-speed interconnects/links must support very high data rates chip-to-chip, board-to-board, and server-to-server communication. Currently, the required aggregate data rates in optical links are in the range of hundreds of gigabits per second (Gbps) (chip-to-chip) and in the range of terabits per second (Tbps) (server-to-server). To support such a high data rate, the bandwidth of the interconnects/links must be increased since the number of available data lanes is not increasing significantly. However, data-rate demands are not constant, and the peak data-rate is not needed all the time. If a link is not adapted for real-time data-rate requirements, then there will be fixed power dissipation, and this will degrade the power-efficiency during periods of sub-maximal link utilization.

A breakdown of the power dissipation of a typical short reach optical link [1] including the transmitter is shown in Figure 1.1. The transmitter dissipates significant amount of power. To improve the energy efficiency of the whole link, the power dissipation of the receiver as well as the transmitter must be decreased at submaximal link utilization. This work focuses on the power reduction at the receiver side.



Figure 1.1. Power breakdown of a typical short reach optical link.

Data communication between servers within a data center is continuously increasing. Datacenters are using parallel optical links with progressively increasing per-lane data rate to meet overall throughput demands. However, the peak data rate is not required by all links all the time [2]. Indeed, links in data-centers are idle up to 90% of the time [3], but non-useful data packets are still being sent to maintain synchronization. Present-day data-centers contain thousands of interconnected servers, and the interconnection network dissipates around 26% of the total datacenter power [4]. As a result, there is a need to reduce the power dissipation of these interconnects when they are idle and reduce the associated cost of cooling data-centers. Links pay a power penalty as they approach the maximum data-rate imposed by the technology in which they are implemented. When a link is not used at its full capacity, packets of unneeded data are still being sent to maintain synchronization. The non-useful data degrades the energy efficiency of the link when we consider energy per information bit. Adapting links to real-time data-rate requirements reduces power dissipation. There are two main approaches for adapting links to real-time data rate requirements: (i) variable data-rate/gear-shifting approach and (ii) burst-mode approach. They have different consequences on the performance of the system. For the burst-mode approach, leaving a link ON permanently would avoid additional time delay between data availability and data transmission, caused by link turn-on and receiver clock synchronization time.

1.2 Objectives/Problems

This work focuses on short-reach links (up to one hundred meters), such as between servers in data-centers, and making them power-efficient during sub-maximal link utilization. A shortlist of the objectives is provided below, followed by their brief explanations.

- Design a rapidly reconfigurable variable-rate optical link with a constant or better power efficiency at a reduced data rate and a constant delay through the front-end to maintain a low bit-error ratio (BER) during the reconfiguration.
- 2. Design a power-efficient burst-mode optical link with an activation time measured in unit intervals (UIs) which is less than the state-of-the-art presented in literature, and shutting down of as much of the circuitry as possible while allowing rapid link synchronization.

Variable-Rate Links

A variable-rate link is a link whose data-rate can be changed to its time-varying demands. This adaptation allows lower power dissipation when operated at a lower data-rate and achieves lower energy per bit [5]. In a variable-rate link, during bandwidth reconfiguration, the delay through the front-end changes. The CDR unit will need to adapt to the phase change (due to the delay change) before an error-free operation is possible. Since CDRs have relatively long time constants, this will incur a latency penalty. The objective is to design a rapidly reconfigurable variable-rate link that has a constant or better energy efficiency per bit at a reduced data-rate and a constant delay through the front-end. The power dissipation is reduced at the reduced data-rate by scaling down the bandwidth of the link.

Burst-Mode Links

In a burst-mode approach, data are transmitted intermittently. Burst-mode links are currently used in point-to-multipoint fiber access systems, such as passive optical networks (PON). The burst-mode approach is a promising way to reduce power dissipation during the idle period and improve the energy efficiency of a link. Two main questions that arise when adapting links to instantaneous data rates are minimum off-state power and minimum link activation time (response time) achievable to maximize the energy efficiency and minimize the BER.

The proposed idea is to power down a link (ideally no power dissipation) during the idle period and power-up it (ideally in no time) during the data transfer. The objective of the work is to develop a burst-mode link based on this proposed method, design its components to achieve link activation time measured in UIs, which is less than the state-of-the-art presented in the literature. The quick activation is done by allowing the complete shutdown all the circuitry except the TIA and a few digital circuits and compare the performance with the present state-of-the-art.

1.3 Claim of Originality

Previous works have presented variable bandwidth front-ends that could be reconfigured quickly. However, changing the bandwidth of the front-end changed the delay through the system. The delay change was never reported in past works. A variable-bandwidth, constant-delay, and power-scalable receiver front-end for a variable-rate optical link is presented. The proposed receiver is capable of rapid switching between two bandwidths of operations. During the bandwidth transition, the receiver front-end was able to maintain constant delay with limited bit errors due to the fixed-delay concept introduced for the first time in the front-end.

Recently, some literature proposed a burst-mode concept for point-to-point communication, such as vertical-cavity surface-emitting laser (VCSEL)-based single-channel optical links. The proposed work makes use of the opportunities arising from the parallel nature of many VCSEL-based links and proposes an energy-efficient burst-mode receiver front-end for multi-channel parallel links for the first time for fast data detection and quick activation of circuitry that is powered off during idle states. The work achieves low power dissipation during idle state, a novel fast data detection with a low bit error rate, and rapid dc-level recovery.

For the first time, the proposed work presents an energy-efficient burst-mode CDR for multi-channel parallel links for fast CDR lock time, exploiting the opportunities resulting from the parallel nature of the VCSEL-based links, as mentioned earlier. The work achieves low-power operation during the idle-state by turning off all circuits in the CDR except the circuit block that generates phase update code. A fast CDR lock time is achieved by updating the code during idlestates using phase update information from an active parallel-channel.

1.4 Publications and Contributions

Peer-Reviewed Journal Articles Directly Related to This Thesis

J1) Abdullah Ibn Abbas, X. Jia, and G. Cowan, "A power proportional, variable-bandwidth and constant-delay front-end for energy-efficient variable-rate optical links," *Submitted to IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Oct. 2020.

Abdullah Ibn Abbas: Designed the overall system, directed the work on optimization of the front-end, planned the block diagram of the circuits, layout, and simulation of the clocking parts, confirmation of the simulations for the front-end, assembly of the top-level chip, testing of the device, and writing of the manuscript.

X. Jia: Worked on the optimization under the guidance of the main author, simulation, and layout of the front-end.

G. Cowan: Supervised project, revision of the manuscript.

J2) Abdullah Ibn Abbas and G. Cowan, "A receiver front-end for VCSEL-based parallel optical links with 49 UI turn-on time," *Submitted to IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Dec. 2020.

Abdullah Ibn Abbas: Designed the overall system, planned the block diagram of the circuits, layout and simulation of the circuits, assembly of the top level chip, measurement plan, testing of the device, and writing of the manuscript.

G. Cowan: Supervised project, revision of the manuscript.

J3) Abdullah Ibn Abbas and G. Cowan, "A fast-locking burst-mode CDR for VCSEL-based parallel optical link receiver," *Submitted to IEEE Access*, Jan. 2021.

Abdullah Ibn Abbas: Designed the overall system, planned the block diagram of the circuits, layout and simulation of the circuits, assembly of the top-level chip, and writing of the manuscript.

G. Cowan: Supervised project, revision of the manuscript.

Peer-Reviewed Journal Papers Not Directly Related To This Thesis

1) C. Williams, D. Abdelrahman, X. Jia, A. Ibn Abbas, O. Liboiron-Ladouceur, G. Cowan, "Reconfiguration in Source-Synchronous Receivers for Short-Reach Parallel Optical Links," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 27, issue 7, p. 1548-1560, Jul. 2019.

Abdullah Ibn Abbas: Took part in initial design of clocking network and ILO, revision of manuscript.

1.4 Thesis Organization

This thesis is organized into six chapters related to the design of power-proportional optical links. Chapter 1 is an introductory chapter. In Chapter 2, some general terms related to a wireline communication system and measurement related topics are reviewed. This chapter also provides a literature review of the relevant work done so far on this topic. Chapter 3 presents the first solution for energy-efficient data communication during sub-maximal link utilization. This solution proposes a rapidly reconfigurable variable-rate receiver without changing the delay through the system during the reconfiguration. Chapter 4 introduces the second solution, called the burst-mode solution, for a link where the data communication is usually in the form of bursts. This chapter provides a fast burst-mode CDR locking technique for a multi-channel parallel optical link. Chapter 5 demonstrates a fast turn-on receiver front-end for the burst-mode multi-channel parallel optical link. Chapter 6 provides the conclusions of the work and presents directions for future research.

Chapter 2 – Background and Literature Review

This chapter presents the necessary background to capture the essence of the proposed work. It starts from the basic circuit and system-level ideas to the measurement related topics. Following this, brief highlights along with the critics of the works done so far, are also presented. Finally, more recent literature works that have led to the proposed work have been discussed in more detail.

2.1 Wireline Communication System



Figure 2.1. Typical communication links: (a) electrical and (b) optical.

In a communication system, the signal is transmitted from one end, called the transmitter (Tx), and received at the other end, called the receiver (Rx), shown in Figure 2.1. The receiver consists of two main circuit blocks: a front-end (FE) and a CDR circuit. The signal suffers from attenuation and distortion all the way from the transmitter to the receiver and therefore needs to be amplified before any further processing. A front-end consists of amplifiers to improve the signal-to-noise ratio (SNR) at the receiver. The whole system consisting of the transmitter, a transmission medium to carry signal, and the receiver is called a link.

The transmission medium may be a coaxial cable, a twisted pair cable, or an optical fiber. If a link has a coaxial cable or a twisted pair cable as a communication medium, it is called an electrical link (Figure 2.1 (a)). In comparison, an optical link has optical fiber (fiber-optic) cable as the communication link (Figure 2.1 (b)).

Therefore, for an optical link, an optical receiver is required. This is shown in Figure 2.2. The two integral parts of an optical receiver front-end are a transimpedance amplifier (TIA) and a post-amplifier (PA), also called a limiting amplifier (LA) or a main amplifier (MA). The focus of this thesis is on an optical link and optical receiver.



Figure 2.2. Conventional optical front-end architecture.

In optical communication, the transmitted signal is in the form of light generated from a laser source. The signal then travels through the optical transmission medium, which is an optical fiber cable, and reaches the receiver end, where it needs to be converted from optical to electrical signal. A photodiode, which is a photodetector, converts an optical signal to an electrical signal. The output electrical signal, generated by the photodiode, is then fed into the receiver front-end, where it is amplified for further processing.

2.2 Transimpedance Amplifier (TIA)

A transimpedance amplifier (TIA) converts a current signal into a voltage signal with an amplification. A TIA is the first stage of amplification at the receiver side. The performance of a TIA is evaluated in terms of its gain, bandwidth, output noise, etc. The gain of a TIA at the receiver side should be high enough to reduce the noise contribution from subsequent stages at the input of the receiver but not too high to limit its bandwidth. There is a tradeoff between the gain and the bandwidth of a TIA given by its fixed gain-bandwidth product. The bandwidth should also be sufficiently high to avoid a large amount of inter-symbol interference (ISI) which corrupts the data signal but again not too high to generate a high integrated output noise. The bandwidth of a TIA is usually chosen between 0.5 - 0.7 of the data rate.



Figure 2.3. (a) Inverter-based shunt feedback TIA. (b) TIA with photodiode model at the input and load at the output.

There are several commonly used topologies of a TIA broadly classified into open-loop and feedback categories. The examples of the open-loop category are common-gate (CG) and regulated cascode topologies, whereas common-source shunt feedback and inverter-based shunt feedback topologies fall under the feedback category. The circuit diagram of an inverter-based shunt feedback TIA is shown in Figure 2.3 (a).

The above shunt feedback TIA consists of an inverting voltage amplifier with gain A_o and feedback resistance R_F . If A_o is sufficiently high, then the closed loop transfer function using Figure 2.3 (b) which includes the photodiode model and the load capacitance is given by

$$Z_{TIA} = -\frac{R_F}{\left(\frac{R_F C_{IN}}{A_0\omega_0}\right)S^2 + \frac{1}{A_0}\left(R_F C_{IN} + \frac{1}{\omega_0}\right)S + 1},$$
(2.1)

where C_{IN} is the total input capacitance (including the photodiode capacitance, C_{PD}), C_L is the load capacitance, and ω_0 is the open loop pole of the voltage amplifier. From (2.1), its low frequency gain (dc gain), also called transimpedance gain represented by R_T , and bandwidth is given by

$$Gain\left(R_T\right) \approx -R_F \tag{2.2}$$

$$\omega_{-3dB} \approx -\frac{A_0}{\left(R_F C_{IN} + \frac{1}{\omega_0}\right)}$$
(2.3)

2.3 Post-Amplifier

A post-amplifier amplifies the output voltage signal of the TIA to a sufficiently high voltage level for proper detection by the decision circuit. A Cherry-Hooper inverter-based topology is shown in Figure 2.4. It consists of two cascaded amplifiers. The first stage converts an input voltage to an output current through its transconductance (g_m). The second stage converts an input current to an output voltage with a gain equal to $\sim -R_F$, where R_F is the feedback resistance in the second stage. The requirement of high gain leads to a few stages in the post-amplifier. However, increasing the number of stages increases power dissipation and noise. Therefore, a typical number of stages for a PA is three to four [6].



Figure 2.4. Cherry-Hooper inverter-based post-amplifier.

2.4 VCSEL

A laser source that has become very popular for many applications is vertical-cavity surface-emitting laser (VCSEL), shown in Figure 2.5. This laser emits light or optical beam vertically from its top surface. The electro-optical characteristics of VCSELs offer the ability to modulate at frequencies exceeding 25 Gbps. These are ideal for high-speed communications. VCSELs are used for communication links in data-centers. The basic performance of a laser is characterized by its wavelength, operating voltage, operating current, output power, and slope efficiency. A symbolic representation of a laser is shown in Figure 2.5 (a), and the die photo (an array of four lasers) of a commercial 850-nm wavelength VCSEL (from Finisar) is shown in Figure 2.5 (b).



Figure 2.5. (a) Symbolic representation of a laser. (b) Micrograph of a commercial (850-nm) VCSEL (an array of four lasers) (Finisar).

2.5 Photodiode

At the receiver side, the optical signal is converted into a current signal by a photodetector called a photodiode (PD), which is shown in Figure 2.6. The input of a photodiode is an optical light signal, and the output is an electrical current signal. A photodiode has a structure of a p-n junction or p-i-n semiconductor with reversed bias. When light is incident on it (on the aperture), photons are absorbed in the photodiode, the electric field sweeps them across and makes it look like a current source. The basic performance of a photodiode is characterized by the wavelength of the incident light, aperture diameter, responsivity, bandwidth, and dark current. The current generated by a photodiode is proportional to the optical power, and the proportionality constant is the responsivity (ρ) of the photodiode. A symbolic representation of a photodiode (from Cosemi) is shown in Figure 2.6 (a), and the die photo of a commercial photodiode capable of detecting 850-nm wavelength optical signal is shown in Figure 2.6 (b). The conversion of incident optical power to generated electrical current by the photodiode is related to its responsivity such that

$$i_{PD} = \rho \times P_{optical},\tag{2.4}$$

where the unit of i_{PD} is in ampere (A), $P_{optical}$ is in watt (W), and ρ is in ampere per watt (A/W).


Figure 2.6. (a) Symbolic representation of a photodiode. (b) Micrograph of a commercial (850-nm) photodiode (Cosemi).

2.6 Bit Error Ratio (BER) and Q-factor

In a practical world, noise is always associated with a signal. The noise is the random, unwanted variation or fluctuation that interferes with the signal. The signal quality is measured through signal-to-noise ratio (SNR) which compares the level of the desired signal to the level of noise. It is expressed as the ratio of signal power to the noise power and often expressed in decibels (dB).

$$SNR = \frac{P_S}{P_N}, \qquad (2.5)$$

where P_S is the signal power and P_N is the noise power. In a communication system, due to the presence of noise in the signal, there is a probability that a certain number of bits in the data sequence are misinterpreted and leads to errors in the transmitted data. Suppose out of 10^9 transmitted bits 1 bit is incorrect, then the bit error ratio is $\frac{1}{10^9} = 1 \times 10^{-9}$. Therefore, SNR is directly related to the bit error ratio (also called bit error rate) (BER) in a communication The BER is a major indicator of the quality of the overall system.

Another source for signal distortion that results from high-frequency bandwidth limitation, insufficient low-frequency cutoff caused by AC-coupling or offset compensation loop, etc., is called inter-symbol interference (ISI).

There are two possible signal levels for the communication of binary data – "high" representing "1" and "low" representing "0". Figure 2.7 shows a binary voltage signal. The originally transmitted "high" and "low" signal levels are denoted by V_H and V_L . The signal in Figure 2.7 is corrupted by an additive white Gaussian noise (AWGN). It is assumed that the two levels of the signals are corrupted by two different noise levels having a standard deviation of σ_H and σ_L , respectively (which are also the rms values of the noises). This means there are two signal-to-noise ratios. In order to accurately calculate the probability of bit error, both signal-to-noise ratios should be considered. The two SNRs can be combined, and the minimum SNR required to obtain a specific BER for a given signal is called the Q-factor. The Q-factor is nothing but a function of the SNR which provides a qualitative description of the receiver performance.

To understand the Q-factor quantitatively and establish a relationship with the BER, some analysis is required to perform. It is also important to note that this relation will be obtained considering only the presence of noise in the signal. Any distortion of the signal due to ISI is neglected. Figure 2.7 shows a binary voltage signal corrupted by an additive white Gaussian noise (AWGN). The originally transmitted "high" and "low" signal levels are denoted by V_H and V_L. It is assumed that the two levels of the signals are corrupted by two different noise levels having a standard deviation of σ_H and σ_L , respectively (which are also the rms values of the noises). The probability distributions of the "high" and "low" signal levels in the presence of AWGN are shown in the figure with standard deviations of σ_H and σ_L and means of V_H and V_L. The probability distribution functions are drawn, assuming that the probability of transmitting "high" and "low" signals is not equal, and also the associated noise levels are different.



Figure 2.7. Probability of error for a binary signal.



Figure 2.8. Probability of error for a binary signal with greater separation between the "high" and "low" values.

In Figure 2.7, v(t) is a sampled signal at time t, V_{Sent} is the originally transmitted level of the binary signal, which can take only two values – V_H and V_L representing "1" and "0". The decision circuit compares the sampled value of the signal to a reference level, V_{th} , called the decision threshold. If the sampled value is greater than V_{th} , it interprets that binary "1" is sampled, whereas if the sampled value is less than V_{th} , it interprets that binary "0" is sampled. It is seen from the figure that the tails of the graphs of probability distribution overlap, and the region is shaded. This shaded region represents the probability of misinterpretation of the sampled values of the signal, i.e., the probability of error. The red region represents the probability that the sampled signal is interpreted as "1" while the originally transmitted signal is $V_{Sent} = V_L$. The blue region represents the probability that the sampled signal is interpreted as "0" while the originally transmitted signal is $V_{Sent} = V_H$. Therefore, the total probability of error is given by

$$P[\varepsilon] = P[\mathbf{v}(t) > \mathbf{v}_{th} | \mathbf{v}_{Sent} = \mathbf{v}_L] P[\mathbf{v}_{Sent} = \mathbf{v}_L] + P[\mathbf{v}(t) < \mathbf{v}_{th} | \mathbf{v}_{Sent} = \mathbf{v}_H] P[\mathbf{v}_{Sent} = \mathbf{v}_H].$$
(2.6)

For mark density of half (i.e., the probability of sending "1" and "0" in the signal is 1/2),

$$P[\mathbf{v}_{Sent} = \mathbf{v}_L] = P[\mathbf{v}_{Sent} = \mathbf{v}_H] = \frac{1}{2}.$$
 (2.7)

In that case,

$$P[\varepsilon] = \frac{1}{2} \times P[v(t) > v_{th} | v_{Sent} = v_L] + \frac{1}{2} \times P[v(t) < v_{th} | v_{Sent} = v_H].$$
(2.8)

The terms $P[v(t) > v_{th} | v_{Sent} = v_L]$ and $P[v(t) < v_{th} | v_{Sent} = v_H]$ are the conditional probability and represent the area under the red and blue regions, respectively. Before presenting the mathematical solutions of the terms in the right-hand side of equation (2.6), it is good to

understand them qualitatively. The target is to know the probability of error in making the decision of the sampled signal. The signal has two levels, and the separation is denoted by $V_{open} = (V_H - \sigma_H) - (V_L + \sigma_L)$ in Figure. 2.7. The closer the separation between the two levels, the higher the probability of error in the presence of noise with a given rms value. The higher the separation, the lower the probability of error, provided the rms values of noise remain the same. This can be visualized through Figure 2.8. Therefore, the error is a function of the value of the signal level and the noise level, or in other words, the ratio between the signal level and the noise level. Now proceeding quantitatively, the solutions for each of these conditional probabilities are derived mathematically using the theory of probability, and the final solution is given by a function called the error function (erf). The error function of a Gaussian random variable x is defined as

$$erf(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{u^{2}}{2}\right)} du.$$
 (2.9)

Hence, equation (2.6) can be written as

$$P[\varepsilon] = \frac{1}{2} \times erf\left(\frac{v_{th} - v_L}{\sigma_L}\right) + \frac{1}{2} \times erf\left(\frac{v_H - v_{th}}{\sigma_H}\right).$$
(2.10)

The arguments of the error functions in equation (2.10) represent the square root of the signal power divided by the square root of the noise power. V_{th} - V_L is the low signal level, and V_H - V_{th} is the high signal level with respect to the decision threshold value. It is interesting to note that the arguments are electrical signals in the form of voltage and can be easily represented in the form of current through division by resistance. These current quantities, when divided by the responsivity of the photodiode (as shown in equation (2.4)), yield optical powers. Therefore, arguments of the error functions represent optical SNRs (SNRO_L for 'low' level, and SNRO_H for 'high' level). Thus, equation (2.10) can be rewritten as follows:

$$P[\varepsilon] = \frac{1}{2} \times erf(SNRO_L) + \frac{1}{2} \times erf(SNRO_H).$$
(2.11)

For the optimum decision threshold level, V_{th-opt} , the probability of bit error is the lowest. Also, at this optimum threshold level, the probability of bit error when a 'high' signal is transmitted, and the probability of bit error when a 'low' signal is transmitted is the same. The equal probability leads to the arguments of the error functions being equal and results in the following definition of the Q-factor:

$$Q \equiv SNRO_L = SNRO_H. \tag{2.12}$$

$$Q = \frac{v_{th} - v_L}{\sigma_L} = \frac{v_H - v_{th}}{\sigma_H}.$$
(2.13)

Finally, the probability of error is given by

$$P[\varepsilon] = erf\left(\frac{v_{th} - v_L}{\sigma_L}\right) = erf\left(\frac{v_H - v_{th}}{\sigma_H}\right) = erf\left(\frac{v_H - v_L}{\sigma_H + \sigma_L}\right).$$
(2.14)

The probability of error, $P[\varepsilon]$, is nothing but the BER. Using equations (2.10) and (2.11), BER can be written as

$$BER = erf(Q) = \int_{Q}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{u^2}{2}\right)} du.$$
 (2.15)

The error function (equation 2.6) is not available in closed form, but for x>3, it can be approximated with high accuracy by

$$erf(Q) \approx \frac{1}{Q\sqrt{2\pi}} e^{-\left(\frac{Q^2}{2}\right)},$$
 (2.16)

Figure 2.9 plots the BER from equation (2.15), and from the plot it is noted that for a bit error rate of 1×10^{-12} , the value of Q is 7.05. This value is frequently used while calculating the sensitivity of a receiver. This means to achieve a BER of 1×10^{-12} the required signal-to-noise ratio is 7.05 in amplitude units. For example, the value of the noise current at the input of a receiver is 3 µA rms, then the required amplitude of the minimum input signal current to achieve a BER of 1×10^{-12} is 21.15 µA (peak).



Figure 2.9. BER vs Q-factor.

The measurement of BER in a laboratory is performed with the help of a bit error rate tester, abbreviated as BERT. The BERT compares the data pattern transmitted by the pattern generator, known as PPG (pulse pattern generator), and the data pattern received by the error detector (ED). It accumulates the number of erroneous bits over time and shows on the screen the accumulated BER from the time of start. There is a 'start' button on a BERT which the user can press to denote the start time of the accumulation of the errors. The displayed BER at any given time is calculated by dividing the number of erroneous bits by the total number of bits received over that period from pressing the 'start' button. Also, the user must wait for a certain period of time to be sure of the target BER.

Suppose the data rate is 10 Gbps and the target BER is 10^{-12} . The bit period corresponding to this data rate is 100 ps. The time span of one bit is 100 ps. Therefore, the time span of 10^{12} bits is 100 s. This means one should wait at least 100 sec (1 min and 40 sec) to decide that the device under test (DUT) results in 1×10^{-12} or less BER. If after waiting for this period of time, the BERT shows greater than 10^{-12} BER, then the DUT fails to meet the targeted BER. However, in practice, one needs to wait 3 to 4 times of this minimum period for the confidence of the specified BER. Table-2.1 shows the minimum time for 1×10^{-12} BER for different data rates.

Table-2.1

Data rate-dependent minimum wait time for10⁻¹² BER

Data rate	Wait time for 10 ⁻¹² BER
1 Gbps	16 min and 40 sec
2 Gbps	8 min and 20 sec
3 Gbps	5 min and 34 sec
4 Gbps	4 min and 10 sec
5 Gbps	3 min 20 sec
6 Gbps	2 min and 47 sec
7 Gbps	2 min and 23 sec
8 Gbps	2 min and 5 sec
9 Gbps	1 min and 52 sec
10 Gbps	1 min 40 sec
>10 Gbps	<1 min and 40 sec

2.7 Extinction Ratio and Optical Modulation Amplitude (OMA)

An extinction ratio is the ratio of two optical power levels of a signal generated by an optical source, e.g., a laser diode (Figure 2.10). The extinction ratio may be expressed as a ratio in linear units or in dB. It is given by

$$r_e = \frac{P_1}{P_0},$$
 (2.17)

where P_1 is the optical power level generated when the signal is high, and P_0 is the power level generated when the signal is low. Figure 2.10 shows the modulated signal (by non-return to zero (NRZ) data) and the power levels corresponding to the high and low condition of the optical signal. The average power will then be given by

$$P_{avg} = \frac{P_0 + P_1}{2}.$$
 (2.18)

The higher the value of the extinction ratio, the lower the DC component (P_{avg}) of the total optical power. The minimum P_{avg} is



Figure 2.10. Non-return to zero (NRZ) data showing the power level.

An extinction ratio can be used to calculate the optical modulation amplitude (OMA). OMA is defined as the difference in optical power between high and low signal levels. Considering the practical scenario, where the quantity available during the measurement is the average optical power (P_{avg}), the OMA is calculated using the P_{avg} , and the extinction ratio, which is given by

$$OMA = 2P_{avg} \left[\frac{r_e - 1}{r_e + 1} \right],$$
 (2.20)

where the extinction ratio is in the absolute form (not in dB or in percentage). OMA represents the optical power and is a useful performance metric of a receiver (expressed in milliwatts or dBm).

$$OMA_{dBm} = 10 \times \log(OMA_{mW}). \tag{2.21}$$

The photodiode converts this optical power to photocurrent (electrical in nature).

2.8 Sensitivity

The sensitivity of an optical receiver front-end at a specified BER is defined as the minimum optical power required at the receiver to achieve that BER. It is the traditional and one of the most widely used specifications of an optical receiver's performance. The better the sensitivity (requiring less optical power), the better is the receiver's performance. The sensitivity of a receiver depends on the total noise generated by each component of the receiver. The sensitivity can also be expressed in terms of photocurrent. One of the ways to calculate the sensitivity of the receiver is to calculate the total noise from each component and referred them to the input of the receiver, called the input-referred noise current. It is straightforward to estimate the rms value of the voltage noise of each component at its output. The output noise (rms value) is referred to the input of the receiver through the division by the total gain from the input of the receiver to that output of the circuit, where the rms noise was estimated. This will result in the rms value of the noise current (because of the TIA presence, the total gain is in the form of V/A). It has been shown earlier that to achieve a specific value of the BER there is a need to maintain a certain

value of the SNR which is given by the Q-factor. Therefore, the following relationship can be written:

$$Q = \frac{i_p}{i_{n,rms}},\tag{2.22}$$

where the i_p is the peak value of the required input signal current and $i_{n,rms}$ is the rms value of the input-referred noise current. The peak-to-peak value of the required signal current (current sensitivity) is found from the following equation:

$$i_{p-p} = 2 \times Q \times i_{n,rms}.$$
(2.23)

For a BER of 10^{-12} , Q=7.05. Therefore, the required value of the peak-to-peak current is 14.1 times that of the rms noise current.

$$i_{p-p} = 14.1 \times i_{n,rms}.$$
 (2.24)

In an optical receiver, often, the reported sensitivity is in the form of optical power. The optical power is represented through OMA which is either in mW or in dBm. The conversion of the current sensitivity to OMA (at sensitivity level) is through the responsivity (ρ) of the photodiode.

$$OMA = \frac{i_{p-p}}{\rho}.$$
(2.25)

$$OMA = \frac{2 \times Q \times i_{n,rms}}{\rho}.$$
(2.26)

Therefore, equation (2.21) can be used to calculate the minimum required optical power for a given BER and input-referred noise current. It is important to note that the measured optical sensitivity will be higher than that calculated by equation (2.21). This is because the Q-factor in the equation is derived considering only the random noise and neglected any distortion in the signal due to ISI.

2.9 Eye Diagram

An eye diagram of a given bit sequence is formed by folding all of the bits into a short interval, as shown in Figure 2.11. The eye diagram provides information regarding a signal in terms of its timing jitter and amplitude variation. The lower the horizontal opening of an eye, the more is the jitter in the signal. The lower the vertical opening of an eye, the more is the amplitude variation in the signal. To obtain an eye diagram of an available random data signal, the total signal length is sliced at a regular interval of time over a given period. This regular interval could be a single or a multiple of a bit period (the time length of a single data bit). Figure 2.11 shows a real 10 Gbps (gigabit per second) data with a bit period of 100 ps (picoseconds) sliced at a regular interval equal to the bit period (T_b) . Also, the choice of start time could result in a circular shift of an eye diagram. Therefore, the start time needs to be adjusted to get the eye diagram centered in the plotted figure. If the regular interval is equal to a bit period, then the diagram will have a shape of a single eye. This is shown in Figure 2.12 (a), plotted for the signal in Figure 2.11. If the interval is equal to two bit periods, shown in Figure 2.12 (b), then the diagram will have two eyes in the eye diagram, which is for the same signal, but the slicing interval is 2T_b. Figure 2.12 (c) and Figure 2.12 (d) show a circular shift in the eye diagrams due to change in the start time of Figure 2.12 (a) and Figure 2.12 (b), respectively. A more practical eye diagram is shown in Figure 2.13 for a large number of data bits plotted over two bits period, and the start time is adjusted to place the eye in the center of the eye diagram. The timing jitter and the amplitude variation is also shown in the figure.



Figure 2.11. (a) Time domain random NRZ data pattern. (b) Slices of the data in (a) at an interval of a bit period to create an eye diagram.



Figure 2.12. Eye diagram of the data shown in Figure 2.10. (a) Eye diagram corresponding to the slicing interval over one bit period or one unit interval (UI) by selecting the start time to position the eye in the center. (b) Eye diagram corresponding to the slicing interval over two bit periods and positioning it in the center with a proper start time. (c) Eye diagram as presented in (a) but with a different start time which leads to a circular shift in the eye. (d) Eye diagram as presented in (b) but with a different start time which leads to a circular shift in the eye.



Figure 2.13. A practical eye diagram with slicing interval over two bit periods and the one eye positioned in the center.

2.10 Transfer Function and Group Delay

The transfer function of a front-end, which includes a TIA and one or more stages of postamplifiers, in the frequency domain is called its frequency response, as shown by equation (2.1) for a TIA. It is a function that evaluates to a complex number. The real part of the complex number is its "magnitude response", and the imaginary part is its phase response. The phase linearity of the front-end is represented by another plot called the "group delay" plot obtained by differentiating the phase response plot with respect to the frequency. By definition, the group delay is given by

$$\tau_g(\omega) = -\frac{d\Phi(\omega)}{d\omega}.$$
(2.27)

The concept of group delay is important as it provides information on the timing delay through a system, and the group delay variation provides knowledge about the timing distortion of a signal caused by the system.

2.11 Phase Noise and Jitter

Phase noise and jitter are closely related. The deviation of zero crossings of a periodic signal from their ideal position in time is called jitter. A time-domain behavior of phase noise is termed as jitter. One of the definitions of jitter is the "period jitter," which is deterministic, and its rms value can easily be calculated by integrating a phase noise spectral density plot. The relationship between the reduction in single-sided phase noise (in dB) and the ratio of increase in power (P) is given by

$$L(f) \propto 10 \log\left(\frac{1}{p}\right). \tag{2.28}$$

2.12 Energy Proportional Links

Links that are not utilized at their full capacity all the time can be made energy proportional by adjusting the links' power dissipation proportional to the amount of data communicated in a time-averaged sense. Depending on the level of link-utilization, links can be adapted for multi-rate or can be made to go from off-state to full data-rate operation known as burst-mode operation.

2.13 Gear-Shifting Link/Variable-Rate Link

A gear-shifting link/variable-rate link is a link in which data-rate is varied depending on the demand for data transfer. In this type of link, there is always link activity (demand for data rate) as shown in Figure 2.14. An example of a power proportional link with constant energy per bit is shown in Figure 2.15.



Figure 2.14. Activity of a gear-shifting link as a function of time.



Figure 2.15. Power dissipation and energy-per-bit of a gear-shifting link [7].

2.14 Burst-Mode Link

In applications where communication links are idle, dissipating full power in the idle states even if no useful data are being transmitted wastes power. This degrades the energy efficiency of links. This is graphically shown in Figure 2.16 (a), where the flat grey area represents a constant power dissipation level, and the hatched area represents the intervals in which the useful data are transmitted. To reduce the power dissipation of these interconnects when they are idle, a promising way is to employ a burst-mode link, where the link is powered off during idle times and powered on with each burst of data. This improves the energy efficiency of the link. The ideal scenario for the power dissipation of a power-proportional burst-mode link is shown in Figure 2.16 (b) where the power is dissipated only when useful data are being transmitted.



Figure 2.16. Conceptual diagram of (a) a conventional link, and (b) a power-proportional burst-mode link, illustrating the power dissipation as a function of link utilization.

If the link is completely off in a powered down state, then the price is paid against a long transient time (response time) to bring the link back to the active state from the state of inactivity. Ideally, zero response time and zero power dissipation in the off state are the desired goals for a burst-mode link.

2.15 Phase Interpolator CDR

A circuit block requires for extracting clock and data retiming is called a clock-and-datarecovery (CDR). The main difference between a phase-locked loop and a CDR is that in a PLL there is a transition in every clock cycle. While in a CDR, there might be or might not be a transition in every clock cycle. A phase interpolator CDR, shown in Figure 2.17, is a dual-loop system with a multi-phase phase-locked loop (PLL) for frequency tracking, and a phase interpolator (PI) for jitter tracking and phase tracking [8]. The phase tracking is required when there is a frequency offset between the transmit side clock and the receive side clock. The PLL used in the CDR provides a frequency nominally equal to the frequency of the data stream. The PI used in the CDR locks to the phase shifts of the data stream. Some advantages of using a phase interpolator CDR are locking speed and digital control for scalability compared to a PLL based CDR. A basic PI CDR has a bang-bang phase detector, an up/down counter with digital control logic, and an analog phase interpolator to interpolate between the available phases of a clock generated by a PLL.



Figure 2.17. Standard phase interpolator CDR block diagram.

2.16 Literature Review

As already mentioned, although there is a demand for high-data-rate links, a given link is not always used at its full capacity. When a link is not required to operate at its full capacity, "required" data are still sent along with extra packets to maintain synchronization. This means that the energy per "required" data bit increases. In a power-proportional link, total energy dissipated by a serial link is proportional to the amount of useful data communicated. Energy proportionality refers to the rate of data communication divided by the power dissipation and is measured in picojoule per bit (pJ/bit). One way of achieving energy-proportionality in such links is through the use of variable-rate links whose power can be scaled proportionally with the data rate. Such links should be capable of rapid changes in data rate. This section will first briefly introduce the literature that addresses link adaptation to the change in the data rate, and then discuss some recent works in detail.

There are two main methods of power reduction in static CMOS circuits. One is by reducing the supply voltage. The other is by lowering the frequency of operation given by the relation $P = CV^2 f$, where P is the power dissipation, C is the load capacitance, V is the supply voltage, and f is the frequency of operation. If static CMOS circuits are used, energy proportionality is expected if the data rate is lowered, and that reducing V at low data rates achieves better than linear scaling of link power dissipation. However, scaling the supply voltage is a slow process and cannot be used for rapid reconfiguration of the link because of the time constant associated with changing the output of a DC-to-DC converter, which provides the supply voltage, is on the order of several microseconds. Data rate scaling for mobile memory interfaces is given in [9], [10]. However, these links are not capable of rapid reconfiguration for various data rates because of supply voltage scaling. Supply-voltage scaling has been adopted for a chip-to-chip interface in [11] and for a memory interface in [9] for energy-efficient bandwidth scaling. A scalable I/O transceiver is presented in [12], and the scaling is obtained by scaling the supply voltage, bias currents, and transmit power with the data rate. Adjusting a link rapidly for various data rates requires its various components such as transmitter driver, clocking circuitry, front-end,

etc., to adapt to this change while maintaining the desired performance. A clock multiplier for a dynamic rate-adjustable interface is presented in [13], where it has been shown that the data rate for a clock forwarded link has been scaled without idle time or bit errors during transitions using fixed frequency ILOs to overcome the frequency relock time. Designing clocking circuitry for an energy-efficient link is challenging because its power must be scaled down during the low data rate mode to save the power for better energy efficiency. A power scalable PLL has been designed in [14] for implementing a CDR suitable for an energy proportional variable data rate link. A complete CDR for a variable I/O-link is presented in [15], where a novel data rate selection logic is used, having a fixed clock operating frequency. However, this work does not support rapid changes in the rate. To achieve improved energy efficiency, the power of the receiver front-end is scaled down by scaling the bandwidth proportionally to the data rate demand. A variablebandwidth front-end with power-scalable capability has been designed in [16], which demonstrates fast response time during the rate transition; however, it did not consider the change in the delay through the system while changing the bandwidth. In conclusion, it is desirable to design a complete receiver considering both the need to rapidly reconfigure the front-end and maintain timing synchronization as the data rate demand changes.

Another method of achieving better energy efficiency in links, which frequently undergo periods of idle time, is by rapidly turning ON or OFF a link. This is referred to as an energy-proportional burst-mode operation. Due to a non-uniform data-rate demand, power can be saved by reducing the average data rate of the link, provided power can be reduced when the link is OFF. Burst-mode techniques are very suitable for many high-speed optical multiaccess network applications such as passive optical networks (PON) [17] which is a point-to-multipoint fiber (optical networks) access system. However, the burst-mode technique used in PON and VCSEL-based links of data-centers are dealt with differently. Burst-mode technique can be employed in PON using time division multiplexing (TDM) [17], where small time slots are provided to all end-users for multiplexing data in time. One end of the link is active all the time and dissipates full power. Although the principal reason for adopting a burst-mode approach for a PON application is not to save power, the power can be saved on the other end which is active only in its time slot. In a VCSEL-link, which is a pointto-point link, this technique will be used to save power on both ends to improve energy efficiency. The common feature of the burst-mode operation applicable to PON and any other energy-proportional link is fast response time, whereas low off-state power is a unique feature of an energy-proportional link. A fast power-on burst-mode transmitter for an energyproportional optical link is presented in [7], whereas [18] demonstrates an optical link transceiver with embedded clock architecture. For a VCSEL based optical transmitter [19], the lasing time of a VCSEL depends on its bias. The smaller the bias, the higher is the time for the laser to start lasing. Hence, there is a tradeoff between the off-state power and the turn on time of the laser. This condition is true for other circuitry also. CDR settling time is a crucial parameter that affects the overall response time of a burst-mode receiver. To achieve fast settling time, [20] presents a hybrid technique that combines the advantage of feedback and feed-forward CDR architectures and recovers the data within 1 unit interval. The fast response time is related to the off-state power. For a complete off-state with no off-state power, the response time will be greater than the state that is not completely off. Other literature works that achieve fast power-on time using various techniques are presented in [21], [22], [23], [24]. A more critical review of related state-of-the-art works is presented in subsequent sections.

There is still an opening and wide scope to develop a multi-channel optical link for energyefficient burst-mode operation.

2.16.1 FE and CDR for Gear-Shifting Link

Previous works that focus on the variable-rate energy-proportional links have been studied and are presented here. For energy-efficient sub-maximal link utilization, the bandwidth of the front-end is reduced during the time of low data rate demand for low power dissipation. [16] presented one circuit that can be reconfigured for operation at different data rates maintaining constant gain. This circuit was designed such that its bandwidth can be reconfigured for different data rates to save power.



Figure 2.18. Variable-rate and power scalable front-end [16].

Figure 2.18 shows the architecture of the front-end. The architecture of the front-end is composed of a TIA with post-amplifier stages and an offset compensation loop. The front-end's bandwidth is adjusted through a tunable resistance bank, and the power is scaled through a binary-

weighted PMOS array. This work highlighted the fact that when the bandwidth of the front-end is scaled down for low data rate operations, the energy efficiency gets better, and the input-referred noise decreases. The decrease in the input-referred noise with the decrease in the data rate will further improve the energy efficiency of the overall link by reducing the required transmitted power. This is presented in Table I, and the energy efficiency is plotted against the data rate in Figure 2.19. However, changing the bandwidth of the front-end will change the delay through the system which was not addressed in [16]. This is an important issue and needs to be considered for a rapidly reconfigurable link to maintain synchronization during the change in the data rate.

Overall receiver from-end performance summary [10]									
Data rate (Gb/s)	1.25	2.5	5	10	15	20			
Gain (dB Ω)	74.1	74.3	74.8	75.1	74.9	74.5			
Bandwidth (GHz)	0.86	1.83	3.78	7.14	10.3	13.1			
Input Referred Noise (pA/ $\sqrt{(Hz)}$)	8.46	10.0	11.6	14.4	15.6	18.0			
Input Referred RMS Noise (μA_{rms})	0.26	0.35	0.63	1.21	1.57	1.89			
Power Dissipation (mW)	0.32	0.78	1.82	5.35	9.41	13.5			
Energy per bit (pJ)	0.26	0.31	0.36	0.54	0.63	0.67			

	T	abl	le-	-2.	.2
--	---	-----	-----	-----	----

Overall receiver front-end performance summary [16]



Figure 2.19. Energy efficiency plot as a function of data rate for the case presented in [16].

The compatibility of this design for a rapid reconfigurable link is shown in Figure 2.20 through a plot of the output voltage of the TIA and the overall receiver front-end. The bias point and the delay time change when the receiver is reconfigured from low data rate operation to high data rate operation. In Figure 2.20, no data were applied to the receiver. The time taken by the circuit to reach a stable operating (95% of the final value) point is referred to as the "response time" here.



Figure 2.20. Response time of the front-end [16].

A CDR circuit is one of the most power-hungry components in a link. Therefore, when it is required to reduce the power of a link at the sub-maximal link utilization to achieve better linkefficiency, the power dissipation of the CDR circuitry must also be reduced. A phaseprogrammable PLL based CDR is presented in [25]. This CDR uses a wide-band phase-locked loop (PLL) to suppress the phase noise from the VCO but is not capable of tuning phase noise and power for a multi-rate operation. This CDR dissipates low power and less die area as compared to other published work [26], [27] for a fixed data rate operation. [15] presented a multi-rate link with a fixed clock rate CDR incorporating a delay-locked loop (DLL). This paper discusses more about the technique of obtaining samples at different rates of link operation but did not address the reduction of the power of the CDR at lower data rates. Although this link is a multi-rate, it is not energy-proportional. The rate selection mechanism is graphically shown in Figure 2.24.



Figure 2.21. Rate selection mechanism. Solid circles represent sample points and hollow circles represent discarded samples [15].

A multi-phase PLL has been designed in [14] to support a rapidly reconfigurable variablerate link CDR. The block diagram of the PLL implementing a VCO-bank is shown in Figure 2.21.



Figure 2.22. Block diagram of PLL with VCO-bank used in [14].

This work takes into consideration the trade-off between phase noise and power dissipation and incorporates a bank of VCOs that are connected and disconnected at high and low data rate operation, respectively. The lower the number of VCOs connected, the lower the power dissipation, as can be seen from (2.28). However, the PLL output jitter will be higher compared to that when a higher number of VCOs is connected. The higher number of VCOs dissipate more power. A decrease of 3 dB and 4.7 dB in open-loop phase noise is expected from the relation given by (2.28) when the number of VCOs connected is increased from 1 to 2 and 3, respectively. This is seen from the open-loop plot in Figure 2.22. The closed-loop phase noise plot shows less reduction in the phase noise compared to the theoretical value because of the presence of other noise sources in the PLL.



Figure 2.23. Open loop and closed loop phase noise plot comparing the number of connected VCOs [14].

The increase in the clock jitter by reducing the power dissipation must be tolerated by the system and should be in harmony with the increase in the bit period of the data at the lower data rate. A variable degree of dynamic phase offset in the output clock, shown in Figure 2.23, called the phase excursion was reported due to activation and deactivation of the VCO-bank and was reduced through the use of variable capacitors at each common node of the connected VCOs. The phase excursion plot for both the compensated and uncompensated situations are plotted in Figure 2.23.



Figure 2.24. Phase excursion plot for dynamic activation of sub-VCO [14].

The technique presented in [14] achieves a power savings of 43% going from high data rate mode to low data rate mode. However, this circuit requires a complex switching sequence, but a simpler circuit is desired which can achieve required savings in power with faster reconfiguration time and lower phase excursion.

A criticism about [15] has been made in [14] that [15] did not discuss switching time when the link goes from one rate of operation to another. [14] addressed this issue in a CDR by adding compensation schemes and then plotting a phase excursion curve during the dynamic sub-VCO activation to estimate the time for stable operation. [16] considered this issue in a variable bandwidth front-end by plotting a response time curve and then determined the time for the reconfiguration.

One important aspect of a variable data rate link that was not considered in all of the above work is the variation in the delay of the data through the system when it is reconfigured for different rates of operation and the degree of maintenance of synchronization with the sampling clock during the reconfiguration. The variable delay and the fixed delay concepts are graphically presented in Figure 2.25. In Figure. 2.25 (a), when the input data-rate is low and the front-end bandwidth setting is also low (LBW), the output data sees a time delay of x. However, when the front-end bandwidth is high (HBW) and the data-rate is still low, the output data sees a lower time delay of y, and the subsequent high-rate output data also undergoes a time delay of y. This instantaneous change in the delay from x to y causes the output data eye to shift relative to the sampling point, leading to a higher BER. On the other hand, in a fixed-delay front-end, the output data shows the same time delay of x for both data-rates even when the bandwidth is changed from low to high, as shown in this Figure 2.25 (b).



Figure 2.25. Data streams showing the input data at a front-end, and (a) the data at the output of a conventional front-end for different bandwidth settings where $x \neq y$ (b) the data at the output of a fixed delay front-end for different bandwidth settings.



Figure 2.26. Data eyes showing the input data of the front-end, and (a) the data at the output of a conventional front-end for different bandwidth settings where x≠y (b) the data at the output of a fixed delay front-end for different bandwidth settings.

A more clearer picture is drawn using an eye diagram concept shown in Figure 2.26. The instantaneous change in the delay will shift the output data eye relative to the sampling clock and

thus cause bit errors during the reconfiguration of the system from low-bandwidth mode to highbandwidth mode. The CDR unit will need to adapt to the phase change before an error-free operation is possible. Since CDRs have relatively long time constants, this will incur a latency penalty. This is graphically presented in Figure 2.26 (a). When the receiver is operating at the lower data-rate (configured to have a lower bandwidth (LBW)), the front-end introduces a time delay of x between the input data and the output data eye, shown in gray. The CDR is assumed to have found optimum sampling locations for edge and data samples as denoted by clock phases S1 and S3, respectively. When the front-end bandwidth is increased (HBW) in anticipation of an imminent increase in the data-rate, the delay of the front-end reduces to y. Subsequent high-rate output data also undergoes a time delay of y. Now, data sample S3 is no longer aligned with the widest vertical eye opening. The reduced eye opening will degrade the BER of the link. This reduction in the eye opening is further intensified in the presence of jitter on both clock and the data.

The fixed-delay concept is graphically shown in Figure 2.26 (b). The output data when the front-end has a low-bandwidth is shown in gray. However, for the fixed-delay front-end, when the front-end is configured to have a high bandwidth, the output data shows the same time delay. The eye diagram with a solid black line represents low rate data when the bandwidth has switched to high mode, whereas the dashed black eye diagram represents subsequent high rate data. In this architecture, we assume that a multiphase oscillator generates clock phases S1 and S3 when receiving low rate data and is capable of generating intermediate clock phases S2 and S4 to sample high rate data without disturbing the position of the already available clock phases (S1 and S3) at low bandwidth mode.

2.16.2 FE and CDR for Burst-Mode Link

Previous literature works dealing with the burst-mode technique in PON and energyproportional links have been highlighted in the literature review section. In a burst-mode transmission, a fast response can improve the efficiency of data transmission by shortening the settling time.

A very significant feature of the proposed point-to-point VCSEL-based link which is different from a point-to-multipoint passive optical network (PON) [17] is the overall improvement in the energy efficiency during idle periods. In PON, the data sent from each end-user (called the optical network unit (ONU)) are in the form of bursts (consisting of useful or non-useful data). The data packets from all ONUs in the network are multiplexed using the time division multiplexing (TDM). PON links are never completely powered down. This is shown in Figure 2.27. These data packets are received at the optical line terminal (OLT). Their amplitude and phase can be quite different on a packet-by-packet basis due to different path loss and difference in each ONU's clock frequency. Therefore, the (OLT) needs to recover the sampling phase of its clock and the average received signal value (dc-level) with each packet (burst) quickly to improve network throughput. However, in a point-to-point VCSEL-based link, the link can be powered down during idle periods and hence increase energy efficiency while achieving fast turn-on time.



Figure 2.27. A conceptual representation of a passive optical network (PON).

The burst-mode technique can be used for point-to-point optical communication, such as in data-centers, where the issues of amplitude and phase variations from burst-to-burst are absent. If a power-proportional burst-mode optical link is put into a sleep mode, the receiver clock loses synchronization, and the average value of the analog output signal changes when links turn off. When a link is turned on as required by a burst of data, its clock will start at an arbitrary phase with respect to the data and take considerable time (~hundreds of UIs) to lock to the appropriate phase for low BER sampling. A typical transient of a burst-mode receiver is shown in Figure 2.28 (a). Turning on a burst-mode receiver involves two tasks. First, the average value of the input signal must be acquired for use as a reference level for the receiver's decision circuit(s). This process is referred to as dc-level recovery. Second, the receiver's sampling clock must be aligned to capture the data through CDR, or in this context, timing recovery. Both these processes are lengthened if the receiver is turned off or put into a low-power state when no input data are received. Lower power during idle states is usually associated with longer turn-on time. Hence, there is a trade-off between the off-state power and the turn-on time, as illustrated graphically in Figure 2.28 (b). This work is focused only on the timing recovery and its rapid activation from a low-power state.



Figure 2.28. (a) Graphical representation of transients in a burst-mode link and (b) conceptual illustration of response times as a function of off-state power.

Recent works on burst-mode solutions for optical networks, wireline electrical, and VCSEL-based optical transceivers have been presented in [18], [19], [23], [28], [29], [30]. The work in [19] focuses on lowering the time for a stable optical output signal from a VCSEL by injecting a high amplitude current pulse to the VCSEL at the beginning of data bursts. This is shown in Figure 2.29. This approach shortens the slow turn-on time of a VCSEL and the recovery time of the reference level for the analog data at the receiver side.



Figure 2.29. Proposed pulse to reduce VCSEL wake-up time [19]. (a) Start preamble with some ns of logic'1's, which will double VCSEL current compared with preamble 101010. (b) A short high-amplitude current pulse is supplied as a wake-up pulse prior to the preamble.

In [23], power-hungry limiting amplifiers (LAs) are replaced with a variable-gain amplifier (VGA), as shown in Figure 2.30, to improve the energy efficiency during the on state. The work uses complicated control loops (Figure 2.30 (a)) to correct the dc-offset current at the input of the TIA and an exhaustive CDR search algorithm to reduce the overall turn-on time of the receiver, which is still in hundreds of UIs. The dc-offset correction feedback loop is formed by the TIA, the VGA, the summer, the TIA calibration logic, and the 6-bit IDAC, as shown in Figure 2.30 (a).



(c)

Figure 2.30. (a) Analog front-end schematic, (b) burst-mode CDR block diagram, and (c) phase-locking dynamics as presented in [23].

The work in [28] reduced the duration of timing recovery by using a complex CDR functionality along with a link protocol presented in Figure 2.31. Both the settling of the dc-level and the CDR locking take place simultaneously in [29] to reduce the overall turn-on time (Figure 2.32). However, all of the solutions discussed above address the turn-on time for a single-channel
burst-mode receiver. These approaches forgo opportunities arising from the parallel nature of many VCSEL-based links.







(b)

Figure 2.31. (a) Link protocol and a burst-mode optical receiver. (b) Power-on time of the receiver [28].



Fast DC correction loop





(b)

Figure 2.32. (a) Burst-mode optical receiver. (b) Transient behavior (simulated) of the receiver [29].

The source synchronous architecture [31] has not been found in the literature being deployed for this type of application in the discussion. Therefore, a typical phase embedded clock CDR is used [32].

2.17. Summary

This chapter provided the frequently used terminology, qualitative and quantitative explanation of certain terms, and basic yet useful concepts regarding an optical communication link. The ideas gained in this chapter will be helpful to understand the following chapters. Following the background topics, state-of-the-art works have been reviewed, and their drawbacks have been discussed.

Chapter 3 – Rapidly Reconfigurable Variable-Rate Optical Link Receiver

This chapter presents a variable-bandwidth front-end for a rapidly reconfigurable variable data-rate and power-proportional optical receiver. The receiver does not detect the data rate change and a rate change signal is required to be transmitted from the transmitter to the receiver for reconfiguration. The design methodology and the measurement of the fabricated design is also presented. The proof-of-concept receiver front-end is capable of operation at 8 Gbps and 4 Gbps data-rates. Implemented in 65 nm CMOS technology, the proposed front-end consists of a shunt-feedback transimpedance amplifier (TIA), a configurable 1-stage to 3-stage post-amplifier and an offset compensation loop. By reconfiguring the number of stages in the post-amplifier, the front-end maintains a near-constant delay when its bandwidth is changed. This allows synchronization to be maintained, with limited bit errors when the target data-rate is switched. The prototype receiver was measured with an optical input at 8 Gbps and 4 Gbps. The overall front-end dissipates 6.12 mW at 8 Gbps (0.76 pJ/bit) and 2.86 mW at 4 Gbps (0.72 pJ/bit). The measurement results confirm the matched delay through the front-end with delay variations within 8 ps.

3.1 Receiver Design

In this work, a receiver has been designed for a variable-rate and power-scalable optical link that can be reconfigured rapidly. The proposed receiver circuit does not detect a change in data rate and some sort of Tx to Rx communication is needed to initiate the reconfiguration. Through reconfiguration, the near constant delay is maintained for discrete values of data rate and these values can be selected at design time. The receiver consists of a power-scalable, variable bandwidth but constant delay front-end and a power-scalable CDR. The conceptual block diagram of the receiver is shown in Figure 3.1.



Figure 3.1. Block diagram of the proposed receiver incorporating the architecture of the variable bandwidth and constant delay front-end.

The receiver operates as a quarter-rate receiver at high data-rate and a half-rate receiver at low data-rate, thus the sampling clock is always at a fixed frequency. This concept is graphically illustrated in Figure 3.2. It is reconfigured by changing the number of post-amplifier stages used and feedback resistors. Since it does not require any change to supply voltage, it can be reconfigured quickly to support high and low data-rates without changing the bias point or perturbing the offset-compensation. The proposed design and brief explanation of each component have been presented which support rapid reconfigurability and power scalability of the proposed variable-rate receiver. The circuits operate at 8 Gbps (high data-rate) and at 4 Gbps (low data-rate) during measurements with optical input.



Figure 3.2. Variable-rate data stream and the sampling clock phases. Shaded clock phases are unavailable during low data rate.

3.2 Power-Scalable, Variable-Rate and Constant-Delay Front-End Architecture

In this work, the proposed front-end consists of an inverter-based shunt-feedback TIA, a three-stage Cherry-Hooper post-amplifier, and an offset compensation circuit similar to [16], [1], and the transistor level circuits are shown in the dotted box in Figure 3.1. The front-end switches between two bandwidths to support high and low data-rates. The offset compensation (OC) circuit used in the proposed front-end, which is an inverter-based topology, corrects offset voltages that occur due to transistor mismatch, and cancels the dc-level of the photocurrent by injecting a current to the input of the front-end. The series resistance used in the OC can be decreased momentarily by an internally generated switching pulse while switching the bandwidth of the front-end to temporarily reduce the time constant for fast dc-level settling.

When a signal passes through a system, each frequency component of the signal sees a different phase response The overall delay of NRZ data through a front-end depends on the dc value (τ_0) of its group delay and group delay variation across the circuit's bandwidth leads to timing distortion. By definition, the group delay is given as

$$\tau_g(\omega) = -\frac{d\Phi(\omega)}{d\omega} \tag{3.1}$$

where ω is the angular frequency and $\Phi(\omega)$ is the phase response. The transfer function of a shuntfeedback TIA has a second order form, given as

$$Z_{TIA}(s) = -R_T \frac{1}{1 + \frac{S}{\omega_0 Q_0} + \frac{S^2}{\omega_0^2}}$$
(3.2)

where R_T is the dc transimpedance gain, and ω_0 is the open loop pole of the voltage amplifier. Its 3-dB bandwidth with a Butterworth response ($Q_0 = 0.707$) is given by

$$BW_{TIA} = f_o = \frac{\sqrt{2A(A+1)}}{2\pi R_F C_T}$$
(3.3)

where A is the open loop amplifier gain, R_F is the feedback resistor, and C_T is the total input capacitance including the photodiode capacitance. The transfer function of a post-amplifier also has a second order form

$$H_{PA}(s) = \frac{A_p}{1 + \frac{S}{\omega_p Q_p} + \frac{S^2}{\omega_p^2}}$$
(3.4)

and the overall transfer function $(Z_T(s))$ with one stage TIA and *n* stage post-amplifier is

$$Z_T(s) = Z_{TIA}(s) \times H_{PA,1}(s) \times \dots H_{PA,n}(s).$$
(3.5)

The number of poles in (3.5) depends on the order of the polynomial in the denominator. Assuming zeros are all well beyond the bandwidth of the circuit, the overall phase response relevant to the proposed technique is obtained by summing the phase contribution of each pole.

$$\Phi(\omega) = \arg \{Z_T(s)\} = -\sum_{k=1}^m \arctan \frac{\omega - b_k}{a_k}$$
(3.6)

where *m* is the number of poles, a_k and b_k are the real and imaginary part of each pole. Therefore, the group delay is found by differentiating (3.6) and has the form

$$\tau_{g}(\omega) = \sum_{k=1}^{m} \left(\frac{a_{k}}{a_{k}^{2} + (\omega - b_{k})^{2}} \right)$$
(3.7)

If a transfer function is frequency scaled by a factor γ , the a_k and b_k terms increase by γ . Therefore, (3.7) shows that τ_g decreases by $1/\gamma$. The simulation results for the delay through the front-end from [16], consisting of a TIA and a single-stage post-amplifier is presented in Figure 3.3. In the simulation, the input data-rate was kept constant at 4 Gbps and the output data were obtained for the front-end bandwidths of 2.5 GHz and 5 GHz. The simulation results show that the change in the delay through the front-end when the bandwidth is changed from 2.5 GHz to 5 GHz is 50 ps which for 8 Gbps data-rate is 40% of a UI, meaning that the CDR must update the sampling point before the link can operate at 8 Gbps, possibly requiring μ s to ms.



Figure 3.3. (a) Transient (dashed for input and solid for output data) and (b) group-delay plots for the delay through the front-end. Grey and black traces for 2.5 GHz and 5 GHz bandwidth respectively.

As seen from (3.6), the delay through a system depends on the group delay which is not only a function of the bandwidth (ω_{-3dB}) of the system but also depends on the number of stages (*n*) given as

$$\tau_d = \tau_g(\omega) = f(\omega, \omega_{-3dB}, n). \tag{3.8}$$

The delay decreases with the increase in ω_{-3dB} but increases with the increase in the number of stages. Therefore, the delay reduction due to increasing the bandwidth can be compensated for by using an amplifier with more stages. In the proposed work the bandwidth of the front-end has been designed for ~60% of the data rate. However, it is configured to have equal delay across the data-rates.

3.3 Front-End Design Methodology

The design methodology with two important aspects is discussed in this section. The two aspects are: (i) the need for reconfigurability in the front-end and (ii) the required number of stages at low data-rate to maintain a constant delay through the front-end. The methodology can be easily modified should the full data-rate design require additional post-amplifier stages. Suppose the front-end has been designed with a TIA and three post-amplifier stages for the required bandwidth and gain to support full data-rate operation. To make the analysis simple, it is assumed that the bandwidth of each second-order Cherry-Hooper post-amplifier stage (f_P) is equal to the TIA bandwidth ($f_P = f_o$). Therefore, the bandwidth of each stage (f_P) in terms of the overall bandwidth of the front-end (f_{-3dB}) is calculated from [33]

$$f_P = \frac{f_{-3dB}}{\left(2^{\left(\frac{1}{n+1}\right)} - 1\right)^{\frac{1}{4}}},$$
(3.9)

where n is the number of post-amplifier stages. The two poles contributed by each second-order stage are found from the calculated bandwidth of each stage, considering the Butterworth response:

$$p_1, p_2 = \frac{\omega_p}{\sqrt{2}} (1 \pm j). \tag{3.10}$$

The group delay is then found using (3.7) which also includes the contribution of poles from the TIA.

Starting from a TIA and three post-amplifier stages designed to operate at high data-rate, it is possible to lower the bandwidth of this system to support low data-rate operation by increasing the feedback resistors R_a and R_b (shown in Figure 3.1). However, changing the bandwidth changes the delay through the system as seen from the discussion of (3.7). A transistor level simulation is presented in Figure 3.4. In Figure 3.4(a), a scaling factor (β) is used to scale the values of R_a and R_b . The vertical marker A1 at β =1 corresponds to the high data-rate (8 Gbps) design point with the desired bandwidth (5 GHz) and gain for the front-end. The vertical marker B1 corresponds to the low data-rate (4 Gbps) design point for which the bandwidth is 2.5 GHz by increasing the values of R_a and R_b through β . It is evident from the plot that to achieve a low bandwidth for low data-rate operation the system suffers a change of 60 ps in the group delay.

In addition to an increase in group delay, increasing the values of R_a and R_b through β also increases the gain from 107 to 124 dB Ω . This motivates the proposed approach in which fewer post-amplifier stages are used to support lower bandwidth/data rate. This allows the lower bandwidth configuration with only one post-amplifier stage to have the same delay as the higher bandwidth configuration with three stages of post-amplifiers and reduces the front-end's power dissipation.



Figure 3.4. Simulated bandwidth, gain, and the group delay of a front-end against the normalized value of the feedback resistor (R_a) with (a) TIA and three- stage post-amplifier, (b) TIA and a single-stage post-amplifier.

Figure. 3.4 (b) is obtained for the front-end with a TIA and a single-stage post-amplifier. The vertical marker A2 corresponds to the low data-rate design point at higher values of R_a and R_b for which the bandwidth is 2.5 GHz. This design has a delay of 125 ps which is closer to the delay of the design A1 (106 ps) than design B1. A2's gain is also closer to A1's than B1's. By eliminating two post-amplifier stages, design A2 dissipates less power than A1 or B1. These simulations suggest that a combination of varying the number of post-amplifier stages and resistor tuning can allow for amplifiers with a bandwidth tailored to each data rate, while maintaining a relatively constant delay.

The analysis above is explored through transistor-level simulation. Simulation results of front-ends with a TIA and no post-amplifier, one-, two- and three-stage post-amplifier are presented in Figure 3.5 as a guideline for the selection of the number of post-amplifier stages for low-bandwidth operation. The design process starts with by determining values for R_a and R_b that lead to a bandwidth of 5 GHz for a front-end with three post-amplifier stages. These resistor values are denoted with $\beta = 1$ in Figure 3.5. This design has a low-frequency group delay of 106 ps. A 2.5 GHz bandwidth line is plotted to investigate possible design points for low data-rate operation with no post-amplifier, one-, two- and three-stage post-amplifier. From the plot it is seen that the bandwidth plots of all of the designs intersect the 2.5 GHz line, however, with different delays and gains. The resulting values of delay are summarized in Table-3.1. From the plot, it is noted that the design with two-stage post-amplifier can achieve equal delay to the three-stage design, but with unnecessarily large bandwidth for 4 Gbps operation, and dissipates more power than the design with only one post-amplifier. However, this configuration could be used for 6.5 Gbps data rate operation with 60% bandwidth design. The designs with one-stage post-amplifier have lower delay differences with the 5 GHz bandwidth design.



Figure 3.5. Circuit level simulation results of bandwidth, gain, and group delay for an inverter-based front-end with a different number of post-amplifier stages against the scaling factor of the feedback resistors R_a and R_b .

Table 3.1

	Number of post-amplifier stages			
Group Delays	(n)			
	0	1	2	3
$ au_{HBW,n}$ (ps)	_	_	_	106
$ au_{LBW,n}$ (ps)	89	125	147	166
$\Delta \tau_n = \tau_{HBW,3} - \tau_{LBW,n} $ (ps)	-17	19	41	60
$\Delta au_n / au_{HBW,3}$	-0.16	0.18	0.39	0.57

Group delays with different number of post-amplifier stages

With $\beta = 3.9$, the design with a single-stage post-amplifier has the desired bandwidth but 125 - 106 = 19 ps (too much) delay. However, the delay for this design can be tuned to match that of the design with a three-stage post-amplifier operating with 5 GHz bandwidth. This is done by slightly lowering the TIA bandwidth and increasing the bandwidth of the post-amplifier while keeping the overall system's bandwidth constant at 2.5 GHz. The scaling of R_a and R_b through scaling factors 'a' and 'b' is shown in Figure 3.6. Scale factors 'a' and 'b' denotes the relative values of R_a and R_b used in Figure 3.6 compared to the design point 'X' in Figure 3.5. This point is also shown as 'P' in Figure 3.6. This figure shows the bandwidth, gain and group delay of the circuit with a single- stage post-amplifier as 'a' is varied from 0.94 to 1.39 and b is varied from 0.05 to 1.05. The selected design has been found by increasing 'a' from 1 to 1.25 and R_b has been decreased by reducing 'b' from 1 to 0.16 to obtain a design point 'Q' with a delay of 110 ps. Hence the design with one post-amplifier stage is selected. It provides a bandwidth of 2.51 GHz, almost

half of the high-bandwidth design with a delay difference of only 4 ps relative to the design with a three-stage post-amplifier. The delay difference is only 0.032 UI of 8 Gbps data.



Figure 3.6. Skewing of R_a and R_b through scaling factors 'a' and 'b' to adjust the group delay in Figure 3.5 using one post-amplifier stage.

As a conclusion, to incorporate reconfigurability in the design, the front-end is switched between three post-amplifier stages and one post-amplifier stage when the bandwidth is switched from high to low while maintaining a near constant delay. At the same time, the resistors are increased when the bandwidth is switched from high to low settings.

Through reconfiguration, the near constant delay is maintained. The simulation of group delay for a front-end with TIA and one stage post-amplifier having an overall bandwidth of 2.5 GHz (low BW) is presented in Figure 3.7. This is compared against the group delay of a TIA with three stage post-amplifier having an overall bandwidth of 5 GHz (high BW). The comparison shows an almost identical dc value of the group delay and hence proves the underlying concept of this work.



Figure 3.7. Simulated group delay of the designed front-end for low and high bandwidth modes at 1.0 V supply voltage. Grey and black traces for 2.5 GHz and 5 GHz bandwidth, respectively.

In this design, it is assumed that if the front-end is configured for 4 Gbps operation and the data-rate needs to be changed to 8 Gbps, the receiver is reconfigured to the high-bandwidth setting

first and the data-rate is changed shortly afterward. When the data-rate is changed from high to low, the data-rate is changed first and then the bandwidth of the front-end is reduced, in order to save power. This sequence is followed to get the receiver ready to receive higher rate data and avoid stressing the data eye due to ISI.

When configured for a 2.5 GHz bandwidth, the output signal is tapped from the first stage of the post-amplifier whereas during high-bandwidth operation each stage can support less gain, and the output signal is taken from the third stage of the amplifier. The additional amplifier stages change the order of the front-end and maintain the same delay while providing the required gain and bandwidth. Introducing the complementary MOS reconfiguration switches in series with the signal path adds extra parasitic resistance and capacitance that degrade the bandwidth by 6%, observed through extracted simulation. The two post-amplifiers are turned off during lowbandwidth operation through MOS switches (shown in Figure 3.1) and resistors R_a and R_b are increased through switches to achieve a desired value of the gain and bandwidth and compensate for the delay change. The simulated power dissipations at 1.0 V supply voltage for low-bandwidth and high-bandwidth settings are 4.08 mW and 8.72 mW respectively. The change in the current drawn from the power supply during the reconfiguration leads to $L \frac{di}{dt}$ effect. This issue is addressed by adding a decoupling capacitor at the supply voltage node of the FE circuit. Hence, the proposed front-end is suitable for a power-proportional variable-rate link. Reconfiguration can be done rapidly since it is done only by opening and closing switches.

3.4 CDR Architecture

One of the most commonly used CDR architectures for plesiochronous clocking is the dualloop structure [8] shown in Figure 3.8. It consists of a cascade of two loops, namely, a core PLL and a peripheral CDR loop. Multiple phases are generated by the PLL and are used by the phase interpolator to align the recovered clock phases to the mid-point of the data streams. The phase interpolator is controlled by the loop filter (LF) through digital code words. In this work, the dualloop CDR has been chosen because it has the advantage of simpler implementation though it suffers from tightly coupled jitter generation and jitter tolerance parameters. In this work, the core PLL has been replaced by an ILO to reduce the power and complexity of the circuit, and to facilitate on-the-fly reconfiguration of the whole CDR. The reconfiguration is done through a software-controlled switch. The alternate stages in the ILO and other unused circuitry at low datarate mode are turned off to meet the number of required clock edges (for half-rate sampling, the number reduces to half compared to the quarter-rate sampling) and the power budget.

The proposed energy-proportional dual-loop CDR architecture is shown in Figure 3.9 and consists of: (1) eight-/four-stage ring ILO, (2) phase rotator (PR), (3) eight/four data/edge samplers, (4) eight/four 1:2 demuxes, (5) 16/8 bang-bang phase detector (!!PD), (6) 16/8-bit majority voting (MV) circuit, and (7) digital loop filter (LF).



Figure 3.8. Dual-loop CDR.



Figure 3.9. Architecture of the proposed CDR.



Figure 3.10. Phase noise and power reconfigurable eight-stage ILO. Shaded blocks are OFF at low datarate mode.

The eight-stage ILO is shown in Figure 3.10 which provides multi-phase clock. For the ILO's output phases to be equally spaced, its free running frequency is tuned to match the

frequency of the reference clock. This is in contrast to the delay-locked loop (DLL) where the delay is tuned to provide equally spaced multi-phase clock. When configured for high data-rate operation, the eight phases required to clock the data/edge samplers at 2.5 GHz are taken from each alternate stage marked as $\varphi 1$ - $\varphi 8$ and dissipates twice as much power as that at low data-rate mode. At low data-rate operation, the alternate faded stages are turned off, and switches are closed to make the oscillator effectively acts like a four-stage ILO to reduce the power dissipation to half. The frequency of oscillation almost remains unaltered, even though the number of stages reduces to half, because each stage's effective loading doubles when the switches are turned on. Also, the phase noise in the jitter tracking bandwidth of the ILO stays the same. The four phases required to clock the data/edge samplers, therefore, are $\varphi 1$, $\varphi 3$, $\varphi 5$, and $\varphi 7$. This ILO structure has the advantage of easy and on-the-fly reconfiguration compared to [14] for high and low data rate with proportional power dissipation.

The eight ILO phases that feed into the phase rotator (PR) are adjusted in phases to sample the data at their mid-points. The designed PR of Figure 3.11 is composed of eight phase selectors (PS) and four phase interpolators (PI). The phase selectors are controlled by the control bits S1-S8 and W1-W16. Two phase selectors and one phase interpolator in each sub-block select two adjacent phases and interpolate between the two selected phases respectively from the given eight phases. Each sub-block operates differentially, and hence four sub-blocks in the PR are needed to produce eight interpolated output clock phases. The selection of phases and the extent of interpolation between the phases depend on the control bits S1-S8 and W1-W16. The critical part of the design is to select reasonable phases, interpolate between the two appropriate phases, and save power when the receiver changes its state from high data-rate operation to low data-rate operation. In this proposed PR, when the data-rate goes from high to low, there is a reasonable power savings (~1/2 the power) by turning off certain blocks while the overall active blocks still produce the desired interpolated output phases. At high data-rate, eight phases, φ_1 - φ_8 , are available from the ILO and the interpolation between the phases φ_1 - φ_2 , φ_2 - φ_3 , φ_3 - φ_4 , φ_4 - φ_5 , φ_5 - φ_6 , φ_6 - φ_7 , φ_7 - φ_8 , and φ_8 - φ_1 produce eight clock phases P1-P8. However, at low data-rate, only four output clock phases from the PR are required to sample the data by interpolating between φ_1 - φ_3 , φ_3 - φ_5 , φ_5 - φ_7 , and φ_7 - φ_1 because we receive only four phases from the ILO and the other four phases are turned off. To achieve interpolation between the desired phases from the ILO at the high data-rate and at the low data-rate and to benefit from power savings a block of selection multiplexers are used, and unused PS and PI are turned off as shown in grey colour in Figure 3.11. In the low data-rate operation, out of the four PIs, two are turned off and out of the eight PSs, four are turned off which accounts for a theoretical overall power savings of almost 50% except for the constant power dissipation in the block of multiplexers.



Figure 3.11. Proposed reconfigurable phase rotator architecture. Shaded blocks are powered down at low data-rate operation.

The sixteen control bits for the PI (W1-W16) and the eight control bits for the PS (S1-S8) are generated by the LF which consists of: 1) 16-bit Right/Left shift register having the ability to set or reset itself depending on the type of binary bit that overflows and whose direction of

movement depends on the Early/Late signal generated by the MV circuit. The Early/Late signal generation by the MV circuit is explained later in this section. 2) 8-bit Right/Left circular 2-HOT bits counter i.e. two bits are '1' at a time. The direction of rotation of this counter depends on the direction of rotation of the 16-bit counter and is only clocked when the 16-bit counter generates SET or RESET signal. During the normal high data-rate operation, the loop filter dissipates full power and produces the control bits for the PS as S1-S8 (Figure 3.12(a)) whereas in the low datarate operation, it dissipates half the power and the control bits for the PS are S1, S2, S5, and S6. These bits have been typed black in Figure 3.12(b) and corresponds to the active PS in the PR during the low data-rate operation. The functionality of the LF is shown through the block diagram in Figure 3.12. The bit-pattern generation of W is simple. If the direction of movement is towards the left corresponding to the EARLY signal, with '1' pushing-in from the right and "0" pushingout from the left, then corresponding to the LATE signal the movement is towards the right with "0" pushing-in from the left and '1' pushing-out from the right. However, the S-bit generation is little different and is graphically explained in Figure 3.12. In the figure, only one direction of bits jumping sequence is shown and the other direction follows the same pattern. In the LF, the power saving at the low data-rate comes only from the 8-bit Right/Left circular counter and the 16-bit Right/Left shift register dissipates the same level of power at both the data rates.



Figure 3.12. Functional block diagram of the digital loop filter and the S-bit generation sequence for (a) high data rate (10 Gbps) and (b) low data rate (5 Gbps).

The incoming data streams are sampled by the samplers for clock recovery and for data recovery. Each sampler used in this design is composed of a comparator [34] and an S/R latch. At the high data-rate operation, four out of the eight samplers are used for data sampling and four for edge sampling whereas at low data-rate operation, two out of the alternate four samplers that remain active are used for data sampling and two for edge sampling and the other four alternate samplers are turned off to save power. Therefore, half the power is saved while switching from high data-rate to low data-rate operations. The outputs from the samplers are down-sampled by 2

using 1:2 demultiplexers to relieve the speed requirement in the digital circuits that follow the samplers. During the full link capacity operation, 8 demultiplexers are active whereas only 4 demultiplexers are active during the half link capacity operation thereby reducing half the power dissipation. The 16 or 8 down-sampled data/edge pairs, depending on the data rate, are used to solve the Alexander equations [35] in the PD block. The PD block consists of two separate subblocks of 16-bit PD and 8bit PD for high data-rate operation and low data-rate operation, respectively, and as a result, 16 Early & Late or 8 Early & Late signals are generated. The PD section of the CDR also provides half the power savings at the low data rate. The 16/8 Early/Late signals are converted to single Early/Late pulses by the MV circuit appropriate for the LF. The MV circuit is also composed of two separate sub-blocks for 16-bit and 8-bit operation and therefore saves half the power at the low data-rate condition. The MV circuits count the number of *Early* and *Late* signals generated by the !!PD and then performs the comparison. If the number of *Early* signal is in majority then it turns its *Early* output high otherwise *Late* is turned high. In the case of a tie, both are turned low. The *Early/Late* signal thus generated drives the LF to generate the correct sequence of control words for the PR to properly synchronize the clock with the incoming data. Hence, the CDR so designed operates at both the modes of operation with a reasonable amount of power savings ($\sim 30\%$) which makes it a suitable candidate for an energy-efficient link.



Figure 3.13. (a) Optical test setup for BER and eye-diagram measurements. (b) ENEPIG finished test board with wirebonded PD and die. (c) Die photo of the receiver chip in 65 nm CMOS.

3.5 Simulation and Measurement Results

Simulation and measurement results of the FE and the CDR are presented in this section. However, for the CDR, only the measurement for the clock generated by the ILO is presented.

Implemented in 65 nm CMOS technology, the die photograph of the prototype receiver is shown in Figure 3.13 (c). The receiver front-end (FE) including the offset compensation loop occupies an area of 0.0174 mm^2 (116 μ m × 150 μ m), of which the offset compensation capacitor alone takes 0.0021 mm^2 of the area. All inputs and outputs are provided with ESD protection. The chip is packaged and wire-bonded in a 44 pin QFN open cavity plastic package for electrical measurements. High-speed probing pads are used for analog input and output. For optical measurements, a bare die was directly wire-bonded on a high-speed PCB. Outputs were taken through SMA connectors (Figure 3.13 (a)).

The fabricated FE dissipates a power of 6.12 mW at 8 Gbps and 2.86 mW at 4 Gbps from 1 V supply voltage and achieves a power reduction of 3.26 mW (53.3%) when reconfigured for lower data-rate.

The low data-rate optical measurements were performed at 4 Gbps with low-bandwidth setting and high data-rate measurements were performed at 8 Gbps with high-bandwidth setting.

All the measurements were performed optically, except the S-parameters measurement which was performed electrically. The test setup used for optical measurements is shown in Figure 3.13 (a). A continuous-wave laser at 850 nm was modulated, after passing it through a polarization controller (PC), by a polarization sensitive modulator having an extinction ratio of 10 dB. The modulating data was an amplified PRBS 2⁷-1 data sequence from a pulse pattern generator (PPG). A variable optical attenuator (VOA) is used to vary the optical power for obtaining optical sensitivities at various data rates. The modulated optical signal is launched to the commercial GaAs PIN photodiode (PD) having a bandwidth of 20 GHz, responsivity of 0.5 A/W, and a typical capacitance of 100 fF. The cathode of the PD is connected to a 2.0 V supply voltage and the anode is wirebonded to the input of the TIA to provide a reverse bias of ~1.53 V. The analog chip output pad is directly wirebonded to the high-speed PCB trace ending in an SMA connector for connectorized off-board measurement.

3.5.1 S-parameters and Gain

The analog FE was measured electrically for its gain and bandwidth performance using an 8.5-GHz Agilent E5071B vector network analyzer (VNA). The gain and bandwidth for two different modes of operation of the FE were obtained from S-parameter measurements at 1 V supply voltage. The measured S-parameters are plotted in Figure 3.14 (a). In the legend, LM and

HM denote low data-rate mode and high data-rate mode, respectively. The gain and the group delay of the FE is also plotted in the same figure (Figure 3.14 (b) and Figure 3.14 (c), respectively) using measured values of S-parameters.

The high-pass characteristic in the figure is due to the offset compensation loop. The sawtooth notch at around 160 MHz in the gain plot is due to the resonance caused by bond-wire inductance and decoupling capacitance in the supply voltage. This is also reflected in the group delay plot. This leads to group delay variation leading to data-dependent jitter in the output signal. The large variation in the group delay plot with respect to frequency is due to the differentiation operation on phase response (3.1) from a relatively low number of saved S-parameter data points during the measurements. However, the average value of the group delay can be estimated to be approximately 112 ps and 105 ps (from 380 MHz to 5 GHz) at low and high modes of operation, respectively with a delay difference of 7 ps. This matches with the simulation results of 107 ps and 102 ps, respectively with a difference of 5 ps shown in Figure 3.7. The overall performance summary of the FE is given in Table 3.2.



Figure 3.14. Measured at 1.0 V: (a) S-parameters, (b) transimpedance gain and (c) group delay of the FE.

Data rate	8 Gb/s	4 Gb/s
Gain (dBΩ)	58.68	64.17
Bandwidth (GHz)	4.50	2.80
Power dissipation (mW)	6.12	2.86

Table 3.2

a rate	8 Gb/s	4 Gb/s
n (dB Ω)	58.68	64.17

FE performance summary

3.5.2 Matched-Delay through the FE

The measurement for the matched delay was performed using the sampling oscilloscope. Through simulations, the delay through the FE was matched at both modes of operation even in the presence of temperature, process and $\pm 5\%$ supply voltage variation (± 0.05 V) shown in Figure 3.15 (a). The measurement results confirm the matched delay through the FE with delay variations lying within 7.3 ps over $\pm 5\%$ supply voltage variations (Figure 3.15 (b)). Figure 3.15 (c) shows an overlaid eye-diagram for delay mismatch of 5.7 ps (0.046 UI) during optical measurement at two different FE bandwidth settings with 4 Gbps data using a supply voltage of 1.0 V. The delay matching did not rely on any off-chip calibration. If this design were ported to a more advanced technology capable of supporting higher bandwidth and hence higher data rate the delay through the front-end would decrease assuming the same circuit topology. Therefore, we expect mismatches in delay between the high and low bandwidth modes to also decrease, remaining a small fraction of the UI.



Figure 3.15. (a) Simulated delay mismatch due to process variation against temperature (i) and in the presence of ±5% supply voltage variation at 27°C for TT process (ii). (b) Optically measured delay mismatch at supply voltages of 0.95 V (i) and 1.05 V (ii). (c) Optically measured overlaid eye-diagrams at low data-rates showing matched delay through the FE at two bandwidth settings for a supply voltage of

3.5.3 Bit Errors during Mode Transition

The measured analog output data from the FE were taken using a 33-GHz Tektronix realtime oscilloscope DPO73304SX before, during, and after the mode transition and are plotted in Figure 3.16 along with the DC level in the absence of a data pattern at the output of the buffer. Since the measurement did not include an on-chip CDR, the analog data was then post-processed using MATLAB. The data is sampled uniformly by a clock with no change in sampling phase to obtain the sampled bits and the corresponding bit errors. The measurement was performed at only one data rate while reconfiguring the receiver from one mode to another mode. This is due to a lack of availability of equipment that can generate data patterns at two different rates and a proper synchronization signal for reconfiguring the receiver.

The test was performed with both transition directions. The number of bit errors in Figure 3.16 (a) depicts one such instance where the output data at 4 Gbps were sampled by a clock while bandwidth switches from low to high. Figure 3.16 (b) is for the case when the bandwidth switches from high to low. The source of the glitch in Fig. 3.16 (a) is due to the saturation when the additional post-amplifiers are powered on. From the figure, it is observed that the output DC level and hence the output data get settled within 6 ns during the bandwidth transition. The 6 ns settling time will cause a maximum of 24 erroneous bits. The number of incorrect bits counted during the transition is 6 during low to high bandwidth transition and 9 during high to low bandwidth transition. Therefore, the minimum time is 6 ns for which the link should remain idle for error-free operation. Hence, the receiver front-end proves to be useful in a rapidly reconfigurable variable-rate link.



Figure 3.16. Analog output data, output DC voltage, and the number of bit errors during reconfiguration.(a) Transition from low bandwidth to high bandwidth mode. (b) Transition from high bandwidth to low bandwidth mode.

3.5.4 FE Power Dissipation vs Bandwidth

In this measurement, power dissipation as a function of bandwidth by tuning the supply voltage is plotted in Figure 3.17 for two different modes: (4 Gbps and 8 Gbps). From the plot of high data-rate mode (shown by the black trace), one can conclude that it is possible to achieve low power dissipation for 4 Gbps operation by further lowering the supply voltage without reconfiguring the FE. However, lowering the supply voltage to adjust the bandwidth keeps the order of the amplifier's transfer function unchanged thereby increasing the delay through the FE. Also, bandwidth adjustment through supply voltage is a slow process and takes a considerable

amount of time (in the range of micro-seconds) to reach a stable operating point. Therefore, rapidly changing the desired bandwidth without changing the supply voltage or the delay needs reconfigurability of the FE which is achieved by a hardware-controlled switch.



Figure 3.17. Power dissipation as a function of bandwidth for different supply voltages.

3.5.5 Sensitivity vs Data-Rate

To further justify the need for reconfigurability in the FE, input sensitivities were measured at various data rates at both bandwidth settings and are plotted in Figure 3.18. The optical measurements for input current sensitivities were performed using a 12.5-Gbps Agilent N4903B J-BERT and a pico-ammeter. The measured average photodiode current was then converted to peak-to-peak value using the extinction ratio. The FE output signal was amplified by an external amplifier to overcome the digital sensitivity of the error detector. From the measured current-sensitivity plots against the data rate, it is observed that the receiver has a sensitivity of 52 μ A_{p-p} at 4 Gbps in low bandwidth mode and 85 μ A_{p-p} at 8 Gbps in high bandwidth mode for a PRBS 2⁷-1 data sequence at a BER of 10⁻¹². The poorer sensitivity at 8 Gbps is due to higher integrated

output noise due to higher bandwidth and higher input referred noise density due to its lower value of feedback resistor. It is also observed that the sensitivity is better at 4 Gbps data-rate in low bandwidth mode of the FE compared to the high bandwidth mode with an additional benefit of 53.3% power reduction through reconfiguration. Both power dissipation and sensitivity improve through reconfiguration. The measured eye-diagram of the output of the front-end at the sensitivity level at both modes is presented in Figure 3.19.



Figure 3.18. Measured current sensitivity for different data-rates at low and high FE modes with optical input.



Figure 3.19. Optically measured eye-diagrams at 4 Gbps (a) and 8 Gbps (b) at sensitivity level optical input powers.

3.5.6 Clock Performance

The on-chip clock generated using the ILO was measured for its performance. Two measurements are important to show the performance of the designed on-chip clock, (1) clock jitter at low and high data-rate of operations while saving power, and (2) low static-phase deskew and fast dynamic-phase settling of the on-chip clock (for data sampling) during data-rate switching. The rms jitter was calculated through measured phase-noise using Rohde & Schwarz FSQ40 Signal Analyzer. The measured ILO dissipates 2.5 mW and 3.5 mW at low-mode operation and high-mode of operation, respectively.

Figure 3.20 is the measured oscilloscope plot of the locked ILO whereas Figure 3.21 shows the phase-noise plots at two mode of operations. The calculated rms jitter at low data-rate mode of operation (2.17 ps) is almost equal to that at high data-rate mode of operation (2.15 ps) because of injection locking.



Figure 3.20. Measured oscilloscope plot of the injection locked clock. (a) low-mode operation and (b) high-mode operation.


Figure 3.21. Phase noise plots obtained during high and low-mode operations.

The instantaneous clock output was captured during the data-rate switching and is shown in Figure 3.22 (a). This effect is reflected in the measured clock deskew phase (Figure 3.22 (b)) with respect to the reference clock. The disruption in the clock generation during the reconfiguration is due to the addition of alternate inverter stages which lead to unsharing of charges at the alternate nodes (Figure 3.10). However, this disruption will not cause major concern in the packet alignment later in the data path because the clock and data skew reamins small after the rate change which is explained in the following sentences. The clock settling time after the switching is around 12 ns and the clock gets settled with a static phase deskew of 16 ps with respect to the static phase deskew before switching. The 12 ns of settling time corresponds to 60 erroneous bits at 5 Gbps during the data-rate switching.

From the measurement, it was observed that the static-phase clock deskew (16 ps) is in the positive direction and half of this deskew will be compensated by the delay mismatch in the FE

(5.7 ps) which is also in the positive direction and hence leaving behind only 10.3 ps which is very small compared to one UI (200 ps) at 5 Gbps.



Figure 3.22. (a) Measured clock output during mode switching and (b) sudden phase drift of the clock with settling time to achieve final static-phase deskew.

3.6 Comparison of the Variable-Rate FE with Published Works

The front-end in [16] is bandwidth variable but did not consider the change in delay during the bandwidth reconfiguration. The front-ends of [1] and [36] rely on supply voltage scaling for varying the bandwidths. Hence, they are not suitable for rapid reconfiguration. The front-end of [37] is capable of automatic reconfiguration. However, the design is not suitable for rapid switching for two reasons. First, the output operating point will change as the load resistance changes and take a long time to settle. Second, the delay through the limiting amplifiers will change. [38] presented an inductor-based front-end in more advanced technology node capable of rapid switching without changing the output dc voltage level. However, this work did not demonstrate constant delay, which therefore necessitates CDR relocking. The presented design is the only power-proportional front-end that maintains a near-constant delay when operated across a range of data rates. This is a critical feature to enable rapidly reconfigurable links.

The power dissipation of the proposed receiver front-end is compared with other published work. The comparison has been summarized in Table 3.3. It is evident from the table that the energy efficiency of this work is comparable with other energy-efficient work in addition to a power reduction of 53.3% while going from high-bandwidth to low-bandwidth operation making it suitable for energy-efficient links.

Table 3.3

FE performance comparison

	JSSC'18	ISSCC'12 [1]		JOSK'16	TCASII'14	JSSC'19	This work	
	[39]	Optimized	Non-optimized	[36]	[37]	[38]		
Data rate (Gb/s)	25	10	22	8.1	25	53	8	4
(Technology)	(65nm)	(90nm)	(90nm)	(65nm)	(40nm)	(28nm)	(65nm)	(65nm)
Measurement (Elec./opt.)	Opt.	Opt.	Opt.	Opt.	Elec.	Opt.	Opt.	Opt.
Gain (dBΩ)	69.4		76	85.1	64		58.6	64.1
Sensitivity $(\mu A_{p\cdot p})$	54		135	56		138	85	52
Measured on-the-fly BW reconFigure	No	No		No	No	No	Yes	
Constant DC operating point with BW reconFigure	No	No		No	Yes	Yes	Yes	
Mechanism of BW adjustment	VDD scaling	VDD and $I_{\rm bias}$ scaling		VDD and feedback resist. scaling	I _{bias} and R _{load} scaling	I _{bias} and R _{load} scaling	No. of stages and feedback resist. scaling	
BW dependent delay	Yes	Yes		Yes	Yes	Yes	No	
Power reduction at low BW	Yes (0.8pJ/bit at 25 Gb/s and 0.55 pJ/bit at 10 Gb/s) (Elec. Meas.)	Yes (0.93pJ/bit at 22 Gb/s and 0.55 pJ/bit at 10 Gb/s)		Yes (8.3 pJ/bit at 8.1 Gb/s and 8.3 pJ/bit at 1.62 Gb/s)	Yes (4.12 pJ/bit at 25 Gb/s and 16.33 pJ/bit at 3 Gb/s)	Yes (0.65 pJ/bit at 53 Gb/s and 0.67 pJ/bit at 27 Gb/s)	Yes (0.76 pJ/bit at 8 Gb/s and 0.72 pJ/bit at 4 Gb/s)	
Power (mW)	30.8	5.5	20.4	67.2 (with CDR)	103	34.6	6.12	2.86
Power efficiency (mW/Gb/s)	1.23	0.55	0.93	8.3	4.12	0.65	0.76	0.72

3.7 Conclusion

A variable-bandwidth, constant-delay, and power-scalable receiver front-end for a variable-rate optical link is presented. As demonstrated through measurement, the receiver is

capable of rapid switching between bandwidths that support 4 Gbps and 8 Gbps operation without changing the DC operating point, while maintaining near-constant input-to-output delay. During the bandwidth transition the receiver front-end was able to maintain constant delay with limited bit errors due to the fixed-delay concept introduced in the front-end. It is expected that if this receiver were implemented in a more advanced CMOS technology, higher bandwidth, and hence higher data rate could be supported. Since this scenario would necessarily have higher frequency poles, input-to-output delay of the front-end would decrease proportionally with bandwidth. Therefore, we expect that the residual mismatch in delay seen when the receiver is reconfigured would decrease proportionally with shrinking UI. Also, a reconfigurable and power-scalable CDR suitable for rapidly reconfigurable variable-rate receiver is presented. The measured performance of the ILO designed for the this CDR is presented during its low- and high-mode of operations.

The complete functionality of the designed CDR could not be presented through measurement. The possible reason for the failure of the fabricated CDR circuit is the tail current source of the phase interpolator did not work properly. Other circuit blocks in the CDR were found to be functional.

Chapter 4 – Burst-Mode CDR for VCSEL-Based Parallel Optical Links

This chapter presents a burst-mode clock-and-data-recovery (CDR) system for multichannel vertical-cavity surface-emitting laser (VCSEL)-based non-return-to-zero (NRZ) optical link's quarter-rate receivers. This work utilizes proxy timing recovery for fast turn-on time. The proxy timing recovery scheme takes advantage of correlated data jitter over parallel optical lanes typically deployed in a data-center. The jitter is correlated as the lanes share the same clocking circuitry and a common reference clock at the transmit side. Rapid timing recovery of the burstmode channel is enabled by incrementing/decrementing its phase rotator (PR) control code during idle periods using phase updates from an always-active channel in the link. This timing recovery approach can be extended to multilevel signaling such as PAM-4. For this type of application, the proposed burst-mode CDR needs to be modified to form a baud rate CDR with low jitter performance, and accommodate the non-uniform threshold at the data edge. The burst-mode timing recovery in such cases should be individual for individual channels to avoid jitter in the clock due to multiple transition crossings at the edge of the data eye, and phase offset of the recovered clock with respect to the data. This chapter also presents circuit-design techniques to reduce power dissipation during idle times while still enabling fast turn-on time. Simulated in 65 nm CMOS technology, the proposed CDR dissipated only 19.5 mW per channel while operating at 10 Gbps/ch and 0.58 mW during its idle state. Simulation results are presented for the turn-on time with the proposed technique and compared against the turn-on time of a conventional receiver. The proposed technique allows the CDR to lock within 26 UIs from when it is powered on irrespective of a 1000 ppm frequency offset between the incoming data and the CDR's reference clock. The complete CDR of each channel occupies an area of 0.045 mm². The proposed scheme introduces only 1.3 % of the area and 2.6 % of on-state power overhead while reducing idle-time power dissipation by 97 %.

The work has two features: 1) low-power operation during the idle state is enabled by turning off all circuits in the CDR except the PR control logic block that increments/decrements the PR's control code; 2) a fast CDR lock time is achieved by updating the PR control code (in the PR control logic) during idle times using a phase update signal from an active parallel channel.

4.1 Phase Relationship of Data over Parallel Channels

Parallel links are widely used to meet high aggregate throughput requirements. Each parallel link has an identical channel and transceiver. Transmitted data streams over parallel links are synchronous from lane to lane as they share the same transmitter-side clocking circuitry and a common transmitter-side reference clock. Also, the jitter is mainly correlated over these links but also have some uncorrelated jitter. A possible mismatch exists in the path length among these channels. However, in serial links, the mismatch in length is not a problem in clock recovery since each lane has its own CDR. Figure. 4.1 shows the phase offset between the data at the receivers of two parallel channels relative to the receiver's reference clock in the presence of parts per million (ppm) frequency offset between the data streams and the receiver's reference clock. The data streams have no ppm offset from one lane to another. One complete cycle around the circle represents one reference clock cycle. *M* denotes the number of data UIs to cover one complete cycle, which is inversely proportional to the ppm frequency offset. Due to the ppm offset, the phase offset for data on channels 1 and 2 increases between t = 0 and t = N UI as well as during successive steps of *N* UI. However, the phase offset between channels 1 and 2 remains small due to their

shared Tx clock. In this figure, N is chosen arbitrarily to be M/5, giving 5 N-UI steps for the phase offset of the data channels to advance around the circle. The small phase offset between channels 1 and 2 will be exploited in a fast recovery of the clock of a burst-mode receiver when it turns on. and is different from that proposed in this work.



Figure 4.1. Phase relationship between the data of two channels in the presence of a ppm frequency offset between data streams and the receiver's reference clock.

The phase relationship of the data streams and their recovered clocks of an always-on channel and a burst-mode channel with respect to the receiver's reference clock is shown in Figure 4.2 (a). φ_1 is the phase difference between the receiver's reference clock and the recovered clock of the always-on channel at *t*=0. φ_2 is the phase difference between the receiver's reference clock and the recovered clock of the burst-mode channel at *t*=0. In this case, the PR control code of the burst-mode channel is not updated while the channel is idle. The righthand side of Figure 4.2 (a) shows the phase relationships after an idle period of *M*/2. The phase of the always-on channel's data and recovered clock has advanced by π , reaching $\varphi_1+\pi$. Since the burst-mode channel's PR control code was not updated, its phase remains at φ_2 , but is misaligned to the burst-mode channel's data which has also advanced by π during the idle period. The phase relationship

is repeated in Figure 4.2 (b) where the PR code of the burst-mode channel is updated between bursts using the proposed timing recovery scheme. Even though most of the burst-mode channel's CDR has been powered down between t=0 and t = M/2 UI, its PR code has been updated using the PR updates from the always-on channel. Therefore, it powers up with a phase offset suitable to capture data from the burst-mode channel.

The advantage of the proposed scheme is that the clock of the burst-mode receiver remains aligned with the data stream from the very first stable clock cycle. The phase detector output of the always-on channel is used as a proxy for phase updates during idle intervals. Hence, this method of timing recovery is referred to as "proxy" timing recovery. The "collaborative" timing recovery scheme used in [40] recovers a global clock from the simultaneous participation of all the channels in the link. The proposed approach, however, has some limitations. All the parallel channels cannot be turned off at the same time. At least one channel should remain active to maintain synchronization. Also, each lane cannot operate at independent data rate.



Figure 4.2. (a) Phase relationship of two data streams, their recovered clocks and the receiver's reference clock, shown at the end of a burst and the start of the next burst on the burst-mode channel. Shown when the burst-mode channel's PR code is not updated. (b) The aforementioned phase relationship with the proposed proxy timing recovery scheme where the PR code of the burst-mode channel is updated between bursts.

4.2 Proposed Architecture and Design

Twelve or more parallel optical-fiber channels are frequently deployed in data-centers where each channel has an independent CDR at the receiver side. This is shown in Figure 4.3 (a). A conceptual block diagram of the proposed twelve-channel burst-mode link is shown in



Figure 4.3 (b) which without the connection among CDRs is similar to a conventional architecture (shown in Figure 4.3 (a)).

Figure 4.3. (a) A conventional twelve-channel parallel optical link architecture. (b) Conceptual block diagram of the proposed burst-mode receiver for a twelve-channel parallel optical link showing one always-on channel and eleven burst-mode channels.

In the proposed architecture, it is assumed that even during periods of low link activity, at least one channel is active, maintaining synchronization. The remaining channels are turned on or off independently depending on the per-channel workload. Each channel has two power states: i) idle-state and ii) on-state. An external half-rate reference clock (5 GHz) is provided for all the

CDRs on the receiver side. During the idle-state, each burst-mode receiver receives a 1/4th rate clock from the always-on channel to detect the beginning of a burst of data, and its PR control code is incremented/decremented using phase updates from the always-on channel. To maximize the jitter correlation between the recovered clock on the always-on channel and the data on the burst-mode channels, the always-on channel is located in the middle (shown in Figure 4.3 (b)).

To investigate opportunities to reduce the turn-on time in parallel optical links, a threechannel receiver prototype is designed, as shown in Figure 4.4. It consists of one always-on channel and two burst-mode channels; the latter operates with rapid on/off functionality to improve energy efficiency during idle times. All channels can operate at 10 Gbps. In the idle-state, most of the circuitry of the burst-mode channel's CDR is turned off to reduce power dissipation. Also, the envisioned front-end of the receiver is configured to have reduced power dissipation and bandwidth, which is presented in the next chapter. The proposed link protocol and the circuit details of the building blocks for the proxy timing-recovery CDR are described in the following subsections.



Figure 4.4. A proposed three-channel burst-mode receiver consisting of one always-on channel and one burst-mode channel.

4.2.1 Proposed Burst-Mode Link Protocol

The proposed link protocol is shown in Figure 4.5 (a). When configured in the idle-state, the front-end is a fully operational receiver, but with reduced bandwidth and power dissipation compared to when it operates in the on-state. The sensing of a data burst is marked by the detection of a few positive edges in the incoming 1/5th rate bit stream. This is achieved by a sensing circuit through a continuous sampling process by a 1/4th rate clock available from the always-on channel. To power on a receiver from its idle-state, an alternating sequence of five "1s" and five "0s" (ON-

BITS) starts the preamble sequence. These ON-BITS, captured by a 1/4th rate decision circuit, produce positive edges that switch the receiver from idle-state to on-state. In this prototype, three positive edges confirm a data burst, thus avoiding any false glitches arising from transients that might appear when the receiver is switched from its on-state to idle-state.



Figure 4.5. (a) Proposed link protocol for the burst-mode receiver. (b) Sampling clock and possible incoming data phases relative to the clock. (c) Sampled data indicating positive edge transitions.

The sequence of 5 bits for each "1" and "0" is to guarantee the three positive edges in the sampled data with a 1/4th rate clock available from the always-on channel. The 1/4th rate sampling clock from the always-on channel is aligned with the midpoint of the data on the always-on channel but the incoming data on the burst-mode channel has an arbitrary phase deskew with the sampling clock of the always-on channel, due to a mismatch in the path lengths of the channels, shown in Figure 4.5 (b). The figure shows possible incoming data phases (i-iii) with respect to the clock from the always-on channel. This clock is used only to detect the incoming burst and is not used for sampling the rest of the burst of data on the burst-mode channel once it is activated. The

corresponding sampled bits are shown in Figure 4.5 (c), and the curved arrows denote positive transitions. The first "0" in each row of sampled data represents the last sampled bit during the idle-state. Even if the first clock edge falls on the metastable position of the data, the circuit with the help of the proposed sequence is able to detect three positive edges. All possible combinations for this case are also presented (square box).

The rest of the preamble sequence (DC-PREAMBLE) is a repeated "1100" pattern of 24 UIs during which the output dc-level is recovered. The useful data follow the preamble sequence, and the link is fully operational. Since the PR control code was updated using phase updates from the always-on channel, the CDR in the burst-mode channel starts from the correct phase with respect to its incoming data burst.

At the end of a data burst, receiving 20 consecutive 0s brings the receiver into the idlestate.

4.2.2 Sense Circuit Description

The sense circuit (SENSE CKT), shown in Figure 4.6, senses the incoming burst of data in 20 UIs and generates the ON-OFF signal to toggle the receiver between the idle-state and the onstate. When ON-OFF = 0, the receiver is in the idle-state. The recovered clock from the always-on channel (CLK1) passes to the latch L1 which samples the output of the front-end in anticipation of an incoming burst of data, ultimately producing three rising edges. The DFFs are clocked by the output of the latch. With the three rising edges, the outputs of the three DFFs are 1, resulting in ON-OFF=1, and thus the clock to latch L1 is gated. Following this, the receiver enters the on-state. When the receiver needs to be placed in the idle-state, the "Off Detect" block (clocked by its own recovered clock, CLK2) generates a reset pulse (RST), after sensing twenty consecutive "0s", which resets the three DFFs inside the "Burst Detect" block and turns ON-OFF=0.

The generation of the RST pulse is governed by the following equation where signals A, B, and Y are shown in Figure 4.6:

$$Y = (A \bigoplus B) \cdot B \tag{4.1}$$
$$= \bar{A} \cdot B \ .$$

The "Off Detect" circuit in the proposed work has been optimized based on the available four data sampling latches for a quarter rate receiver.



Figure 4.6. Circuit details of the SENSE CKT incorporating "Off Detect" and "Burst Detect".

4.2.3 CDR Architecture and Description

A phase interpolator (PI)-based CDR [32] is used in this work as shown in Figure 4.7. It consists of a clock synthesizing block and a peripheral loop. The main circuit blocks of a peripheral loop are a phase detector, a loop filter, phase rotator (PR) control logic, and a PR.



Figure 4.7. A typical phase interpolator-based CDR.

The CDRs of the always-on channel and one burst-mode channel are shown in Figure 4.8. The only difference in the CDR of the burst-mode channel is the selection circuits for the UP/DN and TRIG signals denoted by the shaded region. CLK1 and CLK2 are the clocks generated by the CDRs of the always-on channel and burst-mode channel, respectively. *Early* and *Late* signals are generated from data (D1-D4) and edge (E1-E4) samples by the Alexander phase detector (PD) and the majority voting circuits. They are then processed by the loop filter to create an UP/DN signal (1=UP, 0=DN) and a triggering pulse (TRIG) for the PR control logic. The PR control logic produces a 64-bit code for the PR. The PR generates a differential clock signal, whose rotation covers the whole clock period, and drives the "8-phase generator" block. The outputs of the "8-phase generator" then serve as the clock for the data and edge sampling latches.

The two channels of the proposed architecture interact with each other in such a way that the UP/DN and TRIG signals from the always-on channel feed the PR control logic of the burstmode channel through the selection multiplexers when the ON-OFF signal is low (idle-state). Also, the generated clock (CLK1) from the CDR of the always-on channel is supplied to the SENSE CKT of the burst-mode channel through a gating circuit. In the idle-state, the SENSE CKT continuously checks for the start of a data burst with the help of the CLK1.



Figure 4.8. Block diagram of the CDR architecture of the proposed three-channel parallel optical link showing the always-on channel and one burst-mode channel.

The circuit details and the functionality of the loop filter, consisting of a 4-bit bidirectional shift-register and a finite state machine (FSM), are shown in Figure 4.9. The minimum output update rate of the loop filter is 6 clock cycles. The purpose of using a 4-bit shift register is to have a balance between jitter tolerance and false phase steps due to the overall delay of the CDR loop.

CLK2 passes to the loop filter only when *Early* and *Late* signals are unequal, giving a gated clock, GCK [41]. The advantage of clock gating the loop filter is to reduce the dynamic power of the CDR. The TRIG pulse is generated only when both Q1 and Q4 are high.

The 2.5 GHz differential in-phase and quadrature clock signals, generated by a single 5 GHz reference clock (REF_CK) with the help of an I/Q generator, drive a 128-step CML PR. The output of the PR goes to an 8-phase generator. The 45° spaced output clock phases then drive the high-speed (data sampling) latches.



Figure 4.9. Circuit details of the loop filter and the graphical illustration of the generation of UP/DN and TRIG signals.



Figure 4.10. (a) PR circuit with transistor level diagram of one cell. (b) PR control logic generating 64 control bits for the PR. (c) Block diagram of the 8-phase generator incorporating three cascaded ILOs for the generation of eight clock phases with 45° phase spacing.

The implemented CML PR is shown in Figure 4.10 (a) with circuit level details of one cell [42] and total phase steps of 128 in one clock period (32 phase steps per UI). The low complexity architecture has a conventional round constellation and interpolates between four quadrant clock phases (I+, I-, Q+, Q-) whose weights are determined by two current steering DACs: I-BITS and Q-BITS of the PR control logic. These bits, similar to [41], are shown in Figure 4.10 (b). The PR linearity is strongly influenced by summing node bandwidth. As the data rate increases, the linearity performance of the PR degrades to accommodate the data recovery at the inceresed data rate. The PR control logic consists of two cascaded DACs with an inversion at the point of the cascade. Each of the DACs consists of 32-bit bidirectional shift registers. The functionality of the PR control logic is graphically illustrated in Figure 4.10 (b). The PR control code is zero at power on (or while reset), and the increment/decrement of the code by the UP/DN and TRIG signal from

the loop filter is shown in one direction only when UP/DN=1. Similarly, for UP/DN=0, the DACs will be updated in the opposite direction. The 8-phase generator consists of three cascaded injection locked oscillators (ILOs), shown in Figure 4.10 (c).

Each ILO is a four-stage, cross-coupled, pseudo-differential ring oscillator with four differential injection points. For the first ILO, the differential output of the PR is received at one differential injection point, and the other three injection points are disabled. The other two ILOs receive the four differential outputs of the preceding ILO as the injection signals for their four differential injection points. The four differential outputs of the third s ILO form eight clock phases for data and edge sampling with 45° phase separation. The use of three cascaded ILOs provides better phase separation (each with 45°) compared to only two cascaded ILOs.

The CDR of the always-on channel is in operation at all times with full power dissipation. However, during its idle state, the CDR of the burst-mode channel dissipates a small fraction of its on-state power dissipation due to activity in the PR control logic, necessary to keep the PR code updated by the always-on channel. All other digital blocks in the CDR remain inactive and dissipate no power. The I/Q generator, the PR, and the 8-phase generator are turned off through the ON-OFF signal and dissipate only leakage power in the idle-state. All the switching in the proposed circuit is done simultaneously with the generated ON-OFF signal.

4.2.4 CDR Loop Dynamics, Stability and Inter-Lane Tracking

The CDR used in this design consists of a clock synthesis circuit and a peripheral loop. The clock synthesis circuit includes the external reference clock and the I/Q generator, whereas the peripheral loop comprises the PD, loop-filter, PR control logic, PR, and the ILOs. The ILO used in the proposed CDR provides a multiphase clock. The CDR has a first-order loop (Figure 4.11) and hence it is unconditionally stable. The first order loop is considered for its low complexity as compared to the second order loop. With a frequency offset, a first-order loop will lock with a steady-state phase error that is inversely proportional to the loop gain. The disadvantage of a first-order loop compared to a second-order loop is the static phase error in lock due to frequency offset. In Figure 4.11, the value of K_p is 1. The phase resolution, frequency tolerance, and tracking bandwidth of the proposed CDR architecture are considered next.



Figure 4.11. Block diagram of the implemented CDR used for analysis.

The phase resolution achieved by the designed CDR is given by [43]

Phase Resolution (
$$\Delta \Phi$$
) = $\frac{\text{Clock Period }(T_{CK})}{\text{No. of Accumulator States}}$
= $\frac{4}{128}$ UI
= 31.25 mUI, (4.2)

where 1 UI is the bit period of the incoming data, and the number of accumulator states achieved by two 32-bit shift registers (with a cascading arrangement) in PR control logic is 128.

The maximum tolerable frequency error (Δf) between the incoming data and the nominal frequency ($f_{nom} = 2.5 \text{ GHz}$) of the receiver's reference clock is

$$F_{\text{tol, max}} = \frac{\Delta f}{f_{nom}} \approx \frac{\text{Phase Resolution}}{\text{Min. Phase update interval}}$$
$$= \frac{\Delta \Phi}{6 \times T_{CK}}$$
$$= \frac{31.25 \text{ mUI}}{6 \times (4 \times 1 \text{ UI})} = 1302 \text{ ppm.}$$
(4.3)

The factor 6 in the above equation represents the minimum update clock cycles for UP/DN signals out of the loop filter.

The jitter tolerance (JTOL), which is defined as the maximum sinusoidal jitter amplitude A on the incoming data that the designed CDR can tolerate for a given jitter frequency f_j is given by

$$A < \frac{\Delta\Phi}{2\pi} \cdot \frac{1}{f_j} \cdot \frac{f_{nom}}{6}.$$
(4.4)

Clearly, the higher the frequency of jitter, the lower the amplitude of jitter that can be tolerated. Both (4.3) and (4.4) give theoretical values exceeding what is seen in a real implementation. A reasonable choice of the design parameters was made in this work, and the JTOL curve resulting from (4) is comparable with published work [15].

The use of phase detection from one channel in the updating of the PR code in another channel raises the question of whether the correlation of data jitter seen on adjacent channels leads to detrimental differential jitter, because of the difference in path length between channels [31]. A delay mismatch due to path differences between channels could exceed a few UIs over a 100 m interconnect. Considering clock jitter on the first channel given by [31]

$$J_{C1} = J_A \sin\left(2\pi f_j t\right),\tag{4.5}$$

where J_A is the jitter amplitude and f_j is the jitter frequency. The clock jitter on the second channel in the presence of a path length mismatch is then given by

$$J_{C2} = J_A \sin(2\pi f_j (t + mT_b)),$$
(4.6)

where m is the delay mismatch in number of UIs and T_b is the bit period. Therefore, the differential jitter between the two clocks is given by

$$J_D = J_{C1} - J_{C2}. (4.7)$$

This differential jitter is irrelevant when both lanes operate with separate CDRs. However, when the PR control code on the burst-mode channel is updated based on the always-on channel, J_{C1} is imposed on the burst-mode channel's clock whereas its data has J_{C2} . If J_D exceeds the width of the bathtub curve, errors will occur when the burst-mode channel is activated using the always-on channel's PR updates. For a given jitter amplitude and jitter frequency, as the delay mismatch increases, the differential jitter increases and becomes out of phase, resulting in an increased (BER).

The maximum jitter amplitude given by (4. 4) introduces a timing error (J_D), between the clocks of the always-on and burst-mode channels, less than 0.1 UI when the path difference is up to 78 UIs. This path length mismatch is unlikely to occur in a link of 100m or less. Therefore, updating the PR code of the burst-mode channel during its idle-state will provide a clock, accurate to within 0.1 UI at the onset of a data burst.

4.3 Simulation Results

The functionality of the prototype of the proposed burst-mode receiver (Figure 4.8) is presented in this section. The proposed CDR is designed in a 65 nm, 1 V CMOS process. The total power dissipation of the CDR of each channel operating at 10 Gbps is 19.5 mW (1.95 pJ/bit) during the on-state and reduces to 0.58 mW in the idle-state. The power breakdown of the CDR is presented in Table 4.1. The layout of the entire CDR of one channel is shown in Figure 4.12.

CDR state	Digital Ckt	ILOs	PR	Total	
	(Includes Latches)	+			
	+	Clock-path			
	PR Control Logic				
On-state	3.0 mW	10.8 mW	5.7 mW	19.5 mW	
Idle-state	0.5 mW	0.04 mW	0.035 mW	0.575 mW	

Table 4.1

Idle-state 0.5 mW		0.04 mW 0.035 mV		0.575 mW	
		PR LON R LOCH			
Area = 295 um ²	×153 um	PHASE BRATC			

Power breakdown of the CDR

Figure 4.12. Layout of the CDR in 65 nm CMOS process.

The theoretical value of the JTOL of the CDR used in this work is plotted in Figure 4.13. Two different jitter frequencies are selected (4.14 MHz and 8.28 MHz) for simulating the performance of the CDR with added sinusoidal jitter at these frequencies. The maximum amplitude of jitter tolerated by the CDR was noted (limiting the phase offset variation between the clock and the data within 0.3 UI peak-peak). The noted values of the jitter amplitude are presented in the same plot (Figure 4.13).



Figure 4.13. Theoretical and simulated JTOL of the proposed CDR.



Figure 4.14. Phase delay plot of the phase rotator against phase steps.

There are several non-idealities that stress the performance of the CDR, such as the nonlinearity in the PR, process variation and mismatch effects on the phase interpolator, correlated

and uncorrelated jitter. The PR delay with each phase step is plotted and the maximum phase offset of 6 ps from an ideal value is obtained for a given PR codes. This is shown in Figure 4.14.



Figure 4.15. Mismatch variation at phase rotator code with (a) minimum phase delay error, (b) maximum phase delay error.



Figure 4.16. Process variation at phase rotator code with (a) minimum phase delay error, (b) maximum phase delay error.

Local transistor mismatch will lead the phase interpolators used in the CDRs of the alwayson channel and the burst-mode channels to behave differently. The effect of the mismatch on the phase delay for one hundred runs, at the PR codes that result in the minimum and the maximum phase delay error (Figure 4.14), is shown in Figure 4.15. The phase delays are with respect to the ideal values of delays corresponding to those codes. The means and the standard deviations in the two cases are -2.3 ps and 5.68 ps, and 1.3 ps and 5.4 ps, respectively. The effect of global process variation is shown in Figure 4.16 for these PR codes and the corresponding means and the standard deviations are 1.4 ps and 6.8 ps, and 2.1 ps and 8.7 ps, respectively. The deviation, for a given PR codes, from the average value due to process and transistor mismatch will be handled by the proxy timing-recovery as static phase offset. However, the local transistor mismatch between different codes is of concern. In Figure 4.17 (a), the mismatch is plotted against the run number with phase step number 0 and phase step number 16 (50 ps delay between the two different PR codes). The phase variation between these two codes are highly correlated over the run number. In Figure 4.17 (b) the difference between the phase variations is plotted and it is clear that the maximum phase delay difference is only ~3 ps. Therefore, the proxy timing-recovery works well in the presence of local mismatch and achieves lock when the CDR turns on.



Figure 4.17. (a) Phase delay variation as a function of run number with PR codes resulting in phase step number 0 and phase step number 16. (b) Phase delay difference between these two PR codes as a function of run number.

In an ideal situation, there is a maximum allowable frequency offset between the sampling clock and the data for phase tracking. However, in the presence of a correlated jitter, a lower

frequency offset can be tolerated which depends on the amplitude and frequency of the jitter. For an ideal data eye, the maximum allowable phase offset between the sampling clock and the data, resulting from an uncorrelated jitter, is 1 UI_{p-p} for a specified BER. In the case of completely uncorrelated jitter, a differential jitter of 1 UI_{p-p} (and hence 1 UI_{p-p} clock and data phase offset) is due to an uncorrelated jitter amplitude of 0.5 UI_{p-p} at any given jitter frequency. This uncorrelated jitter between the always-on channel and the burst-mode channel has an adverse effect only when the burst-mode CDR turns on from the idle-state. During the active state, each CDR can track this jitter and remains locked.



Figure 4.18. Tolerated uncorrelated jitter by the proxy timing recovery plotted against the idle period with three different jitter periods of 1000 ns, 242 ns, and 121 ns.

If the idle time is longer than the period of the jitter, then the maximum tolerable amplitude of differential jitter is 1 UI_{p-p} . However, if the time period of the jitter is longer than four times the idle period, the CDR can sustain uncorrelated jitter with an amplitude that results in a differential jitter of more than 1 UI_{p-p} . The phase offset between the sampling clock and the data remains within 1 UI_{p-p} due to uncorrelated jitter if the condition that the time period of the jitter is longer

than four times the idle period is maintained (neglecting other non idealities). The tolerance for uncorrelated jitter at jitter periods of 1000 ns, 242 ns, and 121 ns against the idle periods is plotted in Figure4.18. The longer the period of the jitter compared to the idle period, the higher the differential amplitude of the jitter that can be tolerated by the CDR. For example, if the idle period is 40ns, a 4.14 MHz uncorrelated jitter can be tolerated by the proposed proxy timing-recovery with a differential jitter amplitude (considering anticorrelation) as high as 1.16 UI_{p-p}.

In the proposed design, the overall lock time of the CDR depends mostly on the injection locking of the ILO to the injected clock signal. The extracted simulation (C+CC) of the overall locking behavior of the CDR is presented in Figure 4.19, and it takes 1.9 ns to phase lock to the reference clock (shown in grey) within a phase error of 15 ps (0.15 UI).



Figure 4.19. Simulated (extracted) locking behavior of the complete CDR.

The usefulness of the proxy timing-recovery becomes prominent in recovering the timing information quickly in the presence of a long period without data and with a ppm frequency offset. The transient behavior of the receiver is presented in Figure. 4.20 with a continuous data pattern input to the always-on channel and a burst-mode pattern to the burst-mode channel. The simulation starts with a continuous run of pseudorandom binary sequence (PRBS) 27-1 on the always-on channel and a burst-mode pattern on the burst-mode channel (also PRBS 2⁷-1). To take into account the nonlinearity of the PR in the simulation, the input data pattern of the second burstmode channel is phase shifted by $\frac{1}{2}$ UI and its PR codes are forced to have an offset with the always-on channel. This offset is around 16 phase steps ahead that of the always-on channel, such that when the burst-mode channel turns on, its PR will have near maximum phase delay error (due to the phase update signal from the always-on channel to increment the code by 16 phase steps during the idle period of 40 ns). Simulated plots are presented after an initial period of time over which the sampling clock on the always-on channel and the burst-mode channels are aligned to the optimal sampling position. Figure 4.20 and Figure 4.21 represent the case without the proxy timing-recovery scheme and the case with the proxy timing-recovery, respectively. In the absence of a contribution from the always-on channel to update the control code of the PR in the burstmode channels, the 1000 ppm frequency offset causes the data phase to drift away from the reference clock phase between bursts. Thus, when the burst-mode channels turn on following 40 ns without data, their clocks do not lock within the preamble period. A phase difference of ~40 ps exists between the sampling clock and the data (Figure 4.20). The effect of the PR nonlinearity, as expected, does not have a significant contribution in the phase difference at the end of the idle period. However, in the presence of the phase update contribution from the always-on channel, the PR code is updated, and the sampling clocks of the burst-mode channels get locked within the preamble period of 49 UI (Figure 4.21).



Figure 4.20. Analog output data and sampling clock of burst-mode and always-on channels with ON-OFF signal. Enlarged view of areas before and after the idle-state (40 ns) are also presented in this figure and subsequent figures. The enlarged view is for sampling behavior in the presence of 1000 ppm frequency offset between the transmit and receive clock frequency and without proxy timing recovery scheme. Burst-mode channel-2 also includes the nonlinear effect of the PR. Without the use of proxy timing-recovery, the clocks of the burst-mode channel are not locked at the end of the preamble bits.



Figure. 4.21. The simulation is repeated with proxy timing recovery scheme and only the data and the sampling clocks of the burst-mode channels are shown because the data and the sampling clock of the always-on channel always remain in lock condition after certain period of initialization. The figure shows locking within the preamble period of 49 UI in the presence of 1000 ppm frequency offset (burst-mode channel-2 includes the PR nonlinearity).



Fig. 4.22. The simulation is with proxy timing recovery in the presence of 300 ppm frequency offset (burst-mode channel-2 includes the PR nonlinearity) and 4.14 MHz correlated jitter of amplitude 0.5 UIp-p. The clocks of the burst-mode channels are found to be locked successfully within the preamble period.

In Figure 4.22, the simulation includes correlated jitter in addition to the ppm frequency offset. However, due to the presence of 4.14 MHz correlated jitter with amplitude 0.5 UI_{p-p} , the offset frequency has been reduced to 300 ppm. At the end of the idle period, the clocks get locked within the interval of the preamble period.

The presented ~40 ps difference between the sampling clock and the data is due to an inactive period of only 40 ns chosen for simulation feasibility. Although this 20 ps difference may fall within the bathtub curve, a longer period of inactivity will result in a linearly increasing difference, and the sampling phase will fall outside the bathtub curve. This will result in a high BER even after the 49-UI power-on time, leading to a longer CDR lock time. A graphical representation of the calculated phase difference (in time, Δt) between the sampling clock and the data, resulting from different off periods and 1000 ppm frequency offset, is shown in Figure 4.23.



Figure 4.23. Calculated phase difference (Δt) between the data and the sampling clock resulting due to off period.

Using the proposed technique, irrespective of the length of the off period, the clock remains almost aligned with the data when the CDRs of the burst-mode channels turn on. In this design, the timing recovery depends mostly on the locking time of the ILO. Therefore, the technique presented here proves to be effective in fast timing recovery in a burst-mode multi-channel application.

4.4 Comparison of the Burst-Mode CDR with Published Works

This work has been compared to the most relevant burst-mode works as presented in Table 4.2. The proposed CDR, together with the presented sense circuit, takes less than 49 UIs to lock. However, the CDR takes only 26 UIs from the time it is turned on by the ON-OFF signal. In contrast, the clock recovery time presented in [28] and [28] is 352 UI and 463 UI, respectively from the time the CDR is powered on. The work of [29] concentrates on both timing and output DC-level recovery simultaneously and takes 58 UI for the lock. This work achieves a faster lock time and a comparable power dissipation in [44] implemented in the same technology. The proposed work incorporates ILOs. Their lock time is the most significant factor in determining the overall turn-on time of the CDR. Reducing the lock time of the ILO will correspondingly reduce the lock time of the CDR.

Finally, the power overhead of the proposed design (2.6 %) is better compared to [29] (22 %) working at the same data rate. Therefore, the presented work, which has an area overhead of 597 μ m2 (1.3 % of the non-burst-mode area) coming from the "Off Detect" and the "Burst Detect" circuits, and an on-state power overhead of 0.5 mW, is an attractive solution for energy-efficient parallel optical links.

Table 4.2

CDR performance comparison

	ISSCC'18	ISSCC'15	JOCN'18	JSSC'20	This work	
	[28]	[23]	[29]	[44]		
Tashualary	14-nm	32-nm	130-nm	65-nm	65-nm	
Technology	FinFET CMOS	CMOS	CMOS	CMOS	CMOS	
Data rate (Gb/s)	56	25	10	12	10	
Sanca akt	Vac	No	Not	Vac	Yes	
Sense ckt.	I es	INO	shown	1 65		
CDR lock time (UI)	252	463	58	100	26	
(after powered on)	352		(DC+CDR)	108		
	57	110	110	19.2		
On-state power (mW)	(with DFE logic)	(complete Rx)	20.8	(CTLE and AFE excluded)	19.5	
Off-state power (mW)	0.8		<2	0.96	0.58	
Energy efficiency (pJ/bit)	1.02	4.4	2.08	1.6	1.95	
4.5 Conclusion

This work demonstrates a new approach for fast timing recovery, based on a proxy timingrecovery scheme, and power reduction during the idle period in burst-mode parallel optical links. The proposed architecture is similar to a conventional architecture presently implemented in datacenters. In the proposed technique, the burst-mode channel's PR control code is updated by the always-on channel, while it is in its idle-state.

Simulation results demonstrate a fast CDR lock time of 26 UIs, when it goes from an idlestate to its on-state, in the presence of a ppm frequency offset between the clock and the data. The proposed technique has been demonstrated to be useful in multi-channel parallel links, where at least one channel is active all the time. The clock and the phase update signal from the always-on channel help in fast timing recovery of on-off channels when they go from a low-power state to their on-state.

The functionality of the CDR could not be demonstrated through measurement because the ILOs did not draw current from the power supply. The possible reason for this could be the failure to kick start the oscillation at the free running frequency.

Chapter 5 – Burst-Mode Front-End for VCSEL-Based Parallel Optical Links

This chapter presents a fast turn-on front-end for a burst-mode NRZ receiver targeting VCSEL-based parallel optical links operating at 10 Gbps/ch. The front-end bandwidth and power are reconfigurable for energy-efficient burst-mode operation. The circuit-design techniques for fast power-on using a simple approach is presented. Implemented in a 65-nm CMOS technology, the proposed front-end consists of a shunt-feedback transimpedance amplifier (TIA), a configurable 1-stage or 4-stage post-amplifier and an offset compensation loop. By reconfiguring the number of stages in the post-amplifier from four to one, power dissipation and the front-end bandwidth are reduced during periods of link inactivity while still maintaining sufficient gain to allow the detection of an incoming burst. Rapid power-on is achieved using the regeneration of a high-speed latch activated by an available clock from an always-on parallel channel. The presented work is supported by simulation and measurement results with electrical and optical inputs at 10 Gbps. The measurements demonstrate a 4.9 ns turn-on time, corresponding to 49 UIs.

The proposed work targets fast data detection, and quick activation of circuitry that is powered off during idle states. The work has three features: 1) Low-power operation during the idle-state is enabled by turning off three of four stages in the post-amplifier. Low-bit error rate is maintained in the idle-state by increasing the gain of the active circuitry at the cost of reduced bandwidth. 2) A 1/5th rate preamble sequence compatible with the reduced bandwidth is detected using a high-speed latch and an available 1/4th rate clock to produce a power-on signal. Detection

has low bit-error rate. 3) Rapid dc-level recovery using a reconfigurable offset-compensation loop and the inherently fast power-on time of power-gated inverters.



Figure 5.1. A proposed two-channel burst-mode receiver consisting of one always-on channel and one on-off channel.

5.1 Proposed Receiver Architecture

Twelve or more parallel optical-fiber channels are frequently deployed in data-centers where each channel has an independent CDR at the receiver side. However, transmitters (Tx) on parallel lanes share the same Tx clock, giving rise to precise matching of data rate and correlation of data jitter.

To investigate opportunities to reduce turn-on time in parallel optical links (previously shown in chapter 4), a two-channel receiver prototype is designed, as shown in Figure 5.1. It

consists of one always-on channel and one rapid on-off channel. The latter operates with rapid on/off functionality to improve energy efficiency during idle times. Both channels can operate at 10 Gbps. It is assumed that during periods of low link activity, at least one channel is active and will carry useful data in addition to maintaining synchronization. The remaining channels are turned on or off independently depending on the workload and have two different power states: i) idle-state (low-power state) and ii) on-state. In the idle-state, the front-end is configured to have reduced power dissipation and bandwidth, allowing it to detect an incoming burst but with better sensitivity to the preamble sequence than the on-state has, ensuring accurate detection of incoming bursts. The sensing of a data burst is marked by the detection of a few positive edges in the incoming 1/5th rate bit stream. This is achieved by a sensing circuit through a continuous sampling process by a 1/4th rate clock available from the always-on channel. The proposed link protocol and the circuit details of the building blocks for the receiver front-end are described in the following subsections.

5.1.1 Proposed Burst-Mode Link Protocol

The proposed link protocol is shown in Figure 5.2. This is a modified form of the link protocol shown in section xxx. When configured in the low-power state, the front-end is a fully operational receiver, but with a reduced bandwidth and power dissipation compared to when it operates in the on-state. To power-on a receiver from its low-power state, an alternating sequence of five "1s" and five "0s" (ON-BITS) start the preamble sequence. These ON-BITS, captured by a 1/4th rate decision circuit, produce positive edges that switch the receiver from low-power state to on-state. In this prototype, three positive edges confirm a data burst thus avoiding any false glitches arising from transients that might appear when the receiver is switched from its on-state to low-power state.



Figure 5.2. (a) Proposed link protocol for fast turn-on burst-mode FE. (b) Sampling clock and possible incoming data phases relative to the clock. (c) Sampled data indicating positive edge transitions.

The sequence of 5 bits for each "1" and "0" is to reduce the data rate, compatible with the reduced bandwidth, and to guarantee the three positive edges in the sampled data with a ¹/₄ rate clock available from the always-on channel. The incoming data has an arbitrary phase relationship with the sampling clock shown in Figure 5.3 (b). The figure shows possible incoming data phases (i-iii) with respect to the clock. The corresponding sampled bits are shown in Figure 5.3 (c), and curved arrows denote positive transitions. The first "0" in each row of sampled data represents the last sampled bit during the low-power-state. Even if the first clock edge falls on the metastable position of the data, the circuit with the help of the proposed sequence is able to detect three positive edges. All possible combinations for this case are also presented (square box).

The rest of the preamble sequence (DC-PREAMBLE) is a repeated "1100" pattern of 24 UIs, which is different from the initial preamble sequence, during which the output dc-level is recovered. The increased edge density of this pattern accelerates timing recovery carried out by

the CDR (not presented in this paper) while the front-end completes dc-level recovery. In this interval of time the time constant of the receiver's offset compensation loop is greatly reduced.



Figure 5.3. Block diagram of the proposed receiver incorporating the architecture of the variable bandwidth front-end.



Figure 5.4. Circuit details of the SENSE CKT incorporating "transition from 0 to 1 detection".

5.1.2 Front-End Architecture and Description

In this work, the proposed front-end consists of an inverter-based shunt-feedback TIA, a four-stage Cherry-Hooper post-amplifier (PA), and an offset compensation circuit similar to [16] and [1], which is shown in Figure 5.3. When configured for high-bandwidth operation during the on-state, data are taken from the fourth stage of the PA, whereas during low-bandwidth operation in the low-power state, the data stream is tapped from the first stage of the amplifier. The bandwidth of the front-end has been designed for $\sim 60\%$ of the data rate (10 Gbps) [45] during the on-state. The bandwidth is reduced to support 1/5th of the original data rate (low-bandwidth) during the low-power state. PA stages two through four are turned off in the low-power state, and resistors R_a and R_b are increased through switches to achieve a desired value of the gain and bandwidth. The offset compensation (OC) circuit used in the proposed front-end, which is an inverter-based topology, corrects offset voltages that occur due to transistor mismatch, and cancels the dc-level of the photocurrent by injecting a current to the input of the front-end. The series resistance used in the OC can be decreased momentarily by an internally generated switching pulse while turning on or off the front-end (shown in Figure 5.3) to temporarily reduce the time constant. Thus, in both the on-state and low-power state, the voltage FE OUTPUT settles to the same, mid-rail voltage.

A second aspect of dc-level recovery is the low-pass filtering (LPF) of FE OUTPUT that generates OUTPUT DC-LEVEL which serves as a reference voltage for a differential decision circuit. The LPF eliminates any residual offsets not compensated using the OC loop. However, since FE OUTPUT is nominally unchanged from the low-power state to the on-state, LPF settling time can be ignored when considering the overall dc-recovery time.

The sense circuit (SENSE CKT), shown in Figure 5.4, senses the incoming burst of data very fast and generates the ON-OFF signal quickly to toggle the receiver between the low-power state and the on-state. When ON-OFF = 0, the receiver is in the low-power state. The clock from the always-on channel passes to the latch which samples the incoming burst of data, producing three rising edges. One input of the differential latch is driven by the on-chip FE OUTPUT and the other by OUTPUT DC-LEVEL. The DFFs are clocked by the output of the latch. After three rising edges of data, the outputs of the three DFFs are 1, resulting in ON-OFF=1, and thus the clock is gated. The PMOS switch (M3) quickly activates the three PAs. Following this, the receiver enters the on-state, and the subsequent preamble and data bits see a high bandwidth front-end. When the receiver needs to be placed in the low-power state, an internally generated reset pulse (RST) resets all the three DFFs and turns ON-OFF=0. All the switching in the proposed circuit is done simultaneously with the generated ON-OFF signal. The pulse for reducing the resistance in the OC goes high with the transition of the ON-OFF signal and lasts for about 800 ps. The regeneration of a small analog input signal of the Rx using a latch and the activation of the three PAs through a PMOS switch play an important role in the fast turn-on mechanism as compared to the biasing technique used in [28]. Simulated settling behavior of FE OUTPUT during switching, without input data is shown in Figure 5.5. The settling times for both turning on and turning off are in picosecond range.



Figure 5.5. Simulated settling behavior of the FE (without the input data): (a) while turning on and (b) while turning off.



Figure 5.6. (a) Simulated transfer function with low frequency cut-off. (b) Simulated transient with three positive-edge detection turn-on time and DC-level. (c) Shorter turn-on time with one positive edge transition using a modified circuit.

The proposed FE has been simulated to demonstrate the fast turn-on mechanism and its dclevel settling time. The settling time depends on the time constant of the feedback loop, which is controlled by the low cut-off frequency of the AC response of the FE in the on-state. The simulated low cut-off frequency in the on-state is determined to be 24.7 MHz, as shown in Figure 5.6 (a), and results in a time constant of 6.5 ns. However, by decreasing the OC's resistance by a factor of 100 this value is momentarily reduced to 0.3 ns. Although the time constant of the LPF that computes the output dc-level for the latch is 10 ns, its output does not get disturbed much (because of the fast OC loop). Therefore, the output dc-level settles much before the time constant of the LPF. This is evident from Figures. 5.6 (b) and 5.6 (c). Figure 5.6 (b) shows the turn-on behavior with three positive edges detected in an incoming data burst and is ready for useful data within 49 UIs. However, the FE can turn on with only one positive edge detection using a modified design with only one DFF. This results in a turn-on time of only 29 UIs (Figure 5.6 (c)).

5.2 FE Measurements with Electrical Input Signal

Implemented in 65 nm CMOS technology, the die photograph of the prototype receiver is shown in Figure 5.7 (c). The receiver FE including the offset compensation loop occupies an active area of 0.0077 mm² (95 μ m × 81 μ m), of which the offset compensation capacitor alone takes 0.0041 mm² of the area. All inputs and outputs are provided with ESD protection. A fabricated bare die was directly wire-bonded on a high-speed PCB. For electrical measurements, high-speed probing pads are used for analog input and output and the test setup is shown in Figure 5.7 (a).

The fabricated FE without the output buffer dissipates 5.7 mW during the on-state (10 Gbps) and 1.9 mW during the low-power state from a 0.95 V supply voltage. The FE achieves a power reduction of 67% in the low-power-state. The energy efficiency is 0.57 pJ/bit during the on-state.

For S-parameter measurements, a vector network analyzer (VNA) is used while for BER and sensitivity measurements, a bit error rate tester (BERT) and sampling oscilloscope are used. A real-time oscilloscope is also used for the transient measurements which is not shown in the figure explicitly.



Figure 5.7. (a) Electrical test setup for S-parameter, BER and measurements with probed input and output. (b) ENEPIG finished test board with wirebonded PD and die. (c) Die photo of the receiver chip in 65 nm CMOS.

5.2.1 S-parameters and Gain

The analog FE was measured electrically for its gain and bandwidth performance using an 8.5-GHz Agilent E5071B VNA. The gain and bandwidth for two different modes of operation of the FE were obtained from S-parameter measurements at 0.95 V supply voltage and are plotted in Figure 5.8 (a). In the legend, HBW and LBW denote high bandwidth (on-state) and low bandwidth modes (low-power state), respectively. The gain of the FE is also plotted in the same figure (Figure 5.8 (b)) using measured values of S-parameters. The high-pass characteristic in the figure is due to the offset compensation loop. This leads to group delay variation leading to data-dependent jitter in the output signal. The notch at around 180 MHz in the gain plot is due to the resonance caused by bond-wire inductance and decoupling capacitance in the supply voltage. The overall performance summary of the FE is given in Table 5.1.



Figure 5.8. Measured at 0.95 V: (a) S-parameters and (b) transimpedance gain of the FE.

FE Mode	On-state	Low-power-state
Gain (dBΩ)	60	68
Bandwidth (GHz)	5.9	2.70
Power dissipation (mW)	5.7	1.9

FE performance summary

Table 5.1

5.2.2 Analog Input Voltage Sensitivity and Bathtub Curve

The analog input voltage sensitivities for different data rates at less than or equal to 10^{-12} bit error ratio (BER) were obtained electrically using a 12.5-Gbps Agilent N4903B J-BERT biterror-rate tester and are plotted in Figure 5.9 (a). The input voltage sensitivity for a PRBS 2⁷-1 data sequence at a BER of 10^{-12} is 3.8 mV_{p-p} for 10 Gbps with high bandwidth and is 1.43 mV_{p-p} for 2.0 Gbps with low bandwidth settings. The large input voltage required to obtain a BER of 10^{-12} for 6 Gbps at low bandwidth is due to large ISI caused by limited bandwidth for 6 Gbps data. The sensitivity in the low-power state for 2 Gbps is significantly better than that during the onstate at 10 Gbps. This will guarantee detection of incoming bursts. The measured bathtub curves for two different data rates (10 Gbps and 2.0 Gbps) using input voltages of 1.5x the 10⁻¹² sensitivity level are shown in Figure 5.9 (b). This measurement was performed on the analog FE using a 33-GHz Tektronix real-time oscilloscope DPO73304SX. The wider opening in the case of 2.0 Gbps is due to the higher bandwidth-to-data-rate ratio during the low-power-state. The measured eye-diagram of the output of the front-end at the sensitivity level for both bandwidths is presented in Figure 5.10.



Figure 5.9. Measured (a) voltage sensitivity for different data rates at high and low bandwidths of the analog FE with electrical input and (b) bathtub curves at 1.5x sensitivity level electrical inputs.



Figure 5.10. Measured eye-diagrams at (a) 10 Gbps with high bandwidth and (b) 2.0 Gbps low bandwidth at sensitivity level electrical inputs.

5.2.3 Turn-on Time for Burst-mode Operation

The measured analog output from the FE taken with bursts of data using a 33-GHz Tektronix real-time oscilloscope DPO73304SX is shown in Figure 5.11 (a). Each data burst consists of preamble bits followed by a PRBS 2⁷-1 data pattern. The preamble has on-bits and half-rate data of 1100...1100 (Figure 5.11 (b)). The initial on-bits encounter a low bandwidth front-end and its three positive edges switch the receiver from the low-power state to the on state. Subsequent data see a high bandwidth front-end. From the figure, it is evident that the measured settling time of the FE is less than 4.9 ns which translates to <49 UIs of 10 Gbps data. The calculated dc-level is shown with the dashed line and the signal level does not change much over the period of data burst. Thus, the designed front-end turns on quickly and becomes ready to receive the useful data and hence is suitable for energy-efficient burst-mode links.



Figure 5.11. Output transient behavior of the FE with burst of input data (Electrical input). (a) Multiple data burst, (b) Initialization bits followed by PRBS 2⁷-1 pattern in a single burst.



Figure 5.12. Optical test setup for BER and eye-diagram measurements with probed output.

5.3 FE Measurements with Wire-Bonded Photodiode

The optical measurement setup is shown in Figure 5.12. A Finisar transmitter was used for generating the modulated optical signal. The transmitter consists of an uncooled 850 nm VCSEL. The modulated output signal has an extinction ratio of 9.5 dB. The modulating data was a PRBS 2⁷-1 sequence from a pulse pattern generator (PPG). A variable optical attenuator (VOA) varied the optical power for obtaining the sensitivity at various data rates. The optical signal was launched to a GaAs PIN photodiode (PD) having a bandwidth of 20 GHz, responsivity of 0.5 A/W, and a capacitance of 100 fF. The PD was reverse biased by connecting the cathode to a 2.0 V supply voltage and wirebonding the anode to the input of the TIA. The FE analog output pad was probed for the measurements.



Figure 5.13. Measured (a) current sensitivity for different data rates at high and low bandwidths of the analog FE with optical input and (b) bathtub curves at 1.5x sensitivity level optical inputs.



Figure 5.14. Measured eye-diagrams at (a) 10 Gbps with high bandwidth and (b) 2.0 Gbps low bandwidth at sensitivity level optical inputs.

The optical measurements for input sensitivity were performed using a 12.5-Gbps Agilent N4903B J-BERT. The current sensitivity for a 10 Gbps PRBS 2^7 -1 data sequence at a BER of 10^{-12} during the on-state was 62 μ A_{p-p}. The measured current sensitivity at 10 Gbps and at a BER of 10^{-12} was converted to an optical modulation amplitude using the responsivity of the photodiode and

is calculated to be -9.2 dBm. Also, the analog input current sensitivities for different data rates a BER of less than or equal to 10-12 were obtained optically for both states of the FE and are plotted in Figure 5.13 (a). A significant improvement in the sensitivity in the low-power state compared to the on-state is evident from the plot. Similar to the measurement with an electrical signal input, the measured bathtub curves for two different data rates (10 Gbps and 2.0 Gbps) at two different bandwidths using input currents of 1.5x of the 10-12 sensitivity level are shown in Figure 5.13 (b). Again, the wider opening in case of low bandwidth operation with 2.0 Gbps is due to the higher bandwidth to data rate ratio during the low-power state. The measured output eye-diagrams of the front-end at the optical input sensitivity level for both bandwidths are presented in Figure 5.14.

5.4 Comparison of the Burst-Mode FE with Published Works

This work has been compared to the most relevant burst-mode front-ends as presented in Table 5.2. The proposed front-end, together with the presented sense circuit and PMOS switches, takes 25 UIs to turn-on plus fewer than 24 UIs for dc-level acquisition compared to that presented in [28] (32 UIs to turn-on). The momentarily reduced time constant of the OC circuits allows for fast settling. The proposed front-end has the potential to turn on with the first incoming bit with a modified design, where possible glitches during power-down could be avoided by properly gating the sampling clock. Further, the proposed technique does not rely on having an analog offset compensation scheme. Thus, the dc-level acquisition time could be eliminated by using a digitally controlled offset compensation current source along with the connection to the TIA as used in [28]. Then, the turn-on time is reduced to 5 UIs.

Unlike [28], which uses auxiliary sense and biasing circuitry, this work uses conventional optical receiver circuits to detect incoming data. During the low-power state, the front-end is reconfigured to have lower power dissipation and higher gain than that in the on-state. BER

measurements show improved sensitivity in the low-power-state, which ensures that the preamble bits are always properly detected. Details of data detection and activation of the off circuitry were omitted from [29]. At the sensitivity level, the signal at the output of the TIA in [28] would be small, needing analog amplification to detect it. However, in the proposed work, the analog output data require no further analog amplification. Instead, data are amplified quickly by the high-speed latch.

Finally, the energy efficiency of the proposed design is better compared to [28] and [29]. Therefore, the presented work is an attractive solution for energy efficient parallel optical links.

	ISSCC'18 [28]	JOCN'18 [29]	This work
Technology	14-nm FinFET CMOS	130-nm CMOS	65-nm CMOS
Data rate (Gb/s)	56	10	10
Turn-on sense ckt.	Yes	Not shown	Yes
Switch-on time (UI)	32		<25
On-state power (mW)	59 (with clock- path)	12.1	5.7
Off-state power (mW)	7	2	1.9
Energy efficiency (pJ/bit)	1.05	1.21	0.57

Table 5.2

FE Performance comparison

5.5 Conclusion

A fast turn-on technique for burst-mode parallel optical links has been proposed which allows the link to achieve better energy efficiency during low link activity. The front-end bandwidth and power are reconfigurable for energy-efficient burst-mode operation. The bandwidth and hence the power dissipation of the front-end are reduced during idle periods by reconfiguring the number of post-amplifier stages. Rapid power-on is achieved using the regeneration of a highspeed latch activated by an available clock from an always-on parallel channel. The measurements at 10 Gbps demonstrate a 4.9 ns turn-on time, corresponding to 49 UIs. In this thesis, options for power-proportionality in optical communication links have been explored. A solution for a variable data-rate receiver that adapts to real-time data-rate requirements is proposed. This solution reduces power dissipation and provides low energy per bit during low data rate requirements. The proposed front-end for the variable data-rate receiver maintains synchronization during reconfiguration. This is possible because of a fixed front-end delay when it is reconfigured. A power-efficient burst-mode solution for a multi-channel parallel optical link is also proposed. The proposed work presents a CDR which is powered down during idle periods but still maintains a near-desired phase relationship with the data stream when it is powered up. This significantly reduces the lock time of the CDR. A power-efficient front-end for a burst-mode receiver is presented as well, which is rapidly activated from its inactive state by sensing the incoming data stream. The two proposed techniques (variable-rate and burst-mode) are applied separately in the presented two designs.

6.1 Thesis Highlights

First, a variable-bandwidth front-end is presented for a rapidly reconfigurable variable data-rate and power-proportional optical receiver. By reconfiguring the number of stages in the post-amplifier, the front-end maintains a near-constant input-to-output delay when its bandwidth is changed. The design methodology to determine the number of stages in the post-amplifier to maintain the delay is also provided. The measurement results of the proof-of-concept receiver front-end with optical input at 8 Gbps and 4 Gbps are also presented. The measurement results confirm the matched delay through the front-end

during the bandwidth transition with limited bit errors due to the fixed-delay concept introduced in the front-end. This work has been submitted to the journal *IEEE Transactions on Very Large Scale Integration (VLSI)*.

- Second, an energy-efficient burst-mode CDR system is presented for multi-channel parallel optical link receivers for fast lock time. A newly introduced timing recovery scheme called the "proxy timing recovery" scheme takes advantage of correlated data jitter over parallel optical lanes typically deployed in a data-center. This work also presents circuit design techniques to reduce power dissipation during idle times while still enabling fast lock time. Simulation results are presented for the CDR lock time with a 1/4th rate recovered clock while operating at 10 Gbps/ch. The proposed technique allows the CDR to lock very fast irrespective of a frequency offset between the incoming data and the CDR's reference clock. The proposed CDR can also be ported to the electrical links. This work has been submitted to the journal *IEEE Access*.
- Third, a fast turn-on front-end for a burst-mode receiver targeting parallel optical links is presented. The front-end bandwidth and power are reduced by reconfiguring the number of stages in the post-amplifier during periods of link inactivity for energy-efficient burst-mode operation. A simple circuit design approach is presented for fast data detection and quick activation of circuitry that are powered off during idle states. Rapid power-on is achieved using the regeneration of a high-speed latch activated by an available clock from an always-on parallel channel. Measurements with electrical and optical inputs at 10 Gbps/ch demonstrate a fast turn-on time for the receiver. This work has been submitted to the journal *IEEE Transactions on Very Large Scale Integration (VLSI)*.

6.2 Future Work

This section suggests some possible future research directions related to the work done in the thesis.

6.2.1 Variable-Rate Receiver

As energy efficiency becomes more and more critical for circuit designers, it is desired to introduce a receiver whose power can be scaled with the time-varying data-rate demands. We presented one way to achieve this objective by reducing the bandwidth of the front-end in addition to reducing the power dissipation in the CDR. It was pointed out that bandwidth scaling causes a change in the delay that is predictable.

With such a predictable delay change, it is also possible to rapidly shift the CDR lock point by introducing the correct phase shift in the sampling clock without designing a constant delay front-end. This could be done by incrementing/decrementing a certain number of phase steps rapidly in the phase rotator code. This solution is applicable to a digital CDR. However, it will not be trivial with an analog CDR. This alternative of using an analog CDR could be explored as a future work without compromising the power dissipation.

6.2.2 Burst-Mode Multi-Channel Parallel Receiver Architecture and Related Issues

We presented a multi-channel burst-mode architecture where a channel receives a clock and updates its PR control code during idle-state (i.e., low-power state). The clock is provided by the always-on channel which occupies the middle position in the parallel arrangement of the channels. This technique can be scaled to the silicon photonics solutions, such as ring modulator based links. However, for such application there is a fluctuation in the optical power and therefore the DC offset correction needs to be addressed. There exists a design alternative in terms of injecting the clock, UP/DN and TRIG signals to each CDR. If an adjacent channel is in the onstate, the idle-state channel might take the clock and update its PR code by the adjacent channel for better correlation. This is possible by using selection multiplexers. The extent of the advantage by using the adjacent channel over the always-on channel needs to be studied. Further reduction of power can be achieved by sampling the received data on two adjacent channels by a common CDR because of the correlated jitter, and the static phase offset can be compensated by a deskew circuit.

For a robust and fast response burst-mode receiver design, the following points could be explored:

- When a channel is turned on from the idle-state, it might affect the clock phases in the adjacent already on burst-mode channels. The extent to which this switching phenomenon affects the clock phases needs to be carefully studied, and its effect needs to be minimized.
- The challenge of power supply transient can be avoided by using a regulator that has a very fast response to a change in the load.
- The offset compensation loop in the receiver front-end could be replaced by its digital counterparts to avoid the dc-level settling that results from it.

6.2.3 Burst-Mode ILO

As it was mentioned earlier, the overall lock time of the CDR depends mostly on the injection locking of the ILO to the injected clock signal (when it turns on). Therefore, an ILO with a fast locking time (when it powers up) would result in an overall fast timing recovery. A possible

solution for a fast locking ILO is to use a NAND gate for each stage of an oscillator instead of using an inverter. One input of the NAND gate in each stage is connected to the output of the previous stage and the other input is power gated (connected to an on/off signal).

A simple five-stage inverter-based single-ended oscillator (shown in Figure 6.1) is simulated for its lock time with a 2.5 GHz injected signal when it turns on from its off state. The simulation result is plotted in the form of an overlapping signal over the period equal to the time period of the clock. This is shown in Figure 6.2. In Figure 6.2 (a), the overlapped signal is plotted which starts from the time when the oscillator is powered on (30 ns). It is clearly seen that the oscillator is not locked as soon as it is powered on, and not all the clock segments perfectly overlap each other. Figure 6.2 (b) consists of clock segments starting 4.8 ns after the power on time. This figure also suggests that the lock time is more than 4.8 ns. Finally, Figure 6.2 (c) is for 6 ns after the power on, and all the clock segments are perfectly aligned. Hence, the lock time is 6 ns.

Similarly, a five-stage NAND gate-based single-ended oscillator (shown in Figure 6.3) is simulated with the same 2.5 GHz injected clock, dissipating almost the same power as that for the inverter-based design. The simulation results are presented in Figure 6.4. Figures 6.4 (a) starts at 30 ns which is the power on time of the oscillator, whereas Figure 6.4 (b) is for 400 ps after the power on time of the oscillator. Clearly, 400 ps is the lock time which is significantly less than the inverter-based design. The improvement is because the outputs of the NAND gates are in a known state as one input is low. This is leveraging the technique known as the gated VCOs [46]. Therefore, this approach could be extended for a differential inverter for fast locking when it powers up.



Figure 6.1. Inverter-based five-stage injection locked ring oscillator.



Figure 6.2. Simulated overlapped time periods of the clock signal for determining the lock time of the inverter-based ring oscillator (lock time of 6 ns, shown in (c)).



Figure 6.3. NAND gate-based five-stage injection locked ring oscillator.



Figure 6.4. Simulated overlapped time periods of the clock signal for determining the lock time of the NAND gate-based ring oscillator (lock time of 0.4 ns, shown in (b)).

6.2.4 Redesigning of the Variable-Rate and the Burst-Mode CDR

Since the variable-rate CDR and the burst-mode CDR did not demonstrate their functionality through measurement, future work is to re-fabricate the designs and get the CDRs to work. Future work should address all those issues (phase interpolator's tail current source functionality and the start of the oscillation in the cascaded ILOs) that have likely caused the design failure and further improve performance through research.

Injection Locked Oscillator Design for Variable-Rate CDR

A detailed structure of the ILO used in the variable-rate CDR design is shown in Figure A.1. The designed ILO is a cross-coupled, pseudo-differential current-starved eight-stage ring oscillator.



Figure A.1. Eight-stage ring oscillator using a cross-coupled, pseudo-differential current-starved structure.

When the CDR is operating at high data-rate mode, the available clock phases from the ILO are ϕ_1 , ϕ_2 , ϕ_3 , ϕ_4 , ϕ_5 , ϕ_6 , ϕ_7 and ϕ_8 . Out of the eight clock phases, four phases are used for sampling the data and the other four phases are used for sampling the edges (the Rx acts as a quarter-rate receiver). The differential injection points for the ILO are INJ+ and INJ-. However, when the CDR is operating at low data-rate mode, only four clock phases ϕ_1 , ϕ_3 , ϕ_5 and ϕ_7 are used. Two for data sampling and two for edge sampling (the Rx acts as a half-rate receiver). At

low data-rate mode, the alternate stages in the ring are turned off and by closing the switches across each alternate stage, the oscillator effectively acts as a four-stage ring oscillator.

Injection Locked Oscillator Design for Burst-Mode CDR

A detailed structure of the ILO used in the burst-mode CDR design is shown in Figure B.1. The designed ILO is a cross-coupled, pseudo-differential current-starved four-stage ring oscillator.



Figure B.1. Four-stage ring oscillator using a cross-coupled, pseudo-differential current-starved structure.

When the CDR is active, the tail MOSFETS of the main inverter cells and that of the crosscoupled inverters are connected to the injection signals by throwing the switch towards "INJ" (ON-OFF=1 turns them ON). In case of the first ILO block in the 8-phase generator, only the MOSFETS of the first pair of the main inverter cells (producing phases ϕ 1 and ϕ 5) are connected to the injection signals. The remainings are connected to the high voltage level (or VDD). The other two ILO blocks have the connections as shown in the above figure. When the CDR is in the low-power state, the ILOs are turned off by throwing the switches of all the tail MOSFETS towards "GND" (ON-OFF=0 turns them OFF).

On-chip Shift Register Design for Control Bits

For the measurement purpose of a fabricated chip that requires control signals, an on-chip N-bit shift register is designed to provide the required control signals. This is shown in Figure C.1. The figure shows an on-chip shift register whose outputs are connected to the respective pins of the on-chip circuit for the proper functioning of the complete integrated circuit (IC). The shift register has three input pins which are driven by an off-chip microcontroller.



Figure C.1. An N-bit shift register showing its input and output connections.



Figure C.2. The N-bit shift register consisting of small and identical sub-blocks.



Figure C.3. Connection of DFF in each sub-block of the N-bit shift register.

To design an *N*-bit shift register, where *N* is a large value, a practical way of creating it is to first design an *M*-bit shift register, where *M* is a small number (*M*<<*N*). This is shown in Figure C.2. This technique reduces the complexity of the schematics and the layout. Each sub-block is a regular structure. The required number of control bits can be extended by cascading the sub-blocks using a hierarchical design approach (*M*=*N*×*L*; *L* is the number of required sub-blocks).

Each sub-block consists of two rows, each with M number of edge-triggered flip-flops (DFF), shown in Figure C.3. The bottom row of M number of DFFs acts as a serial-input and

parallel-output shift register. In contrast, the top row of *M* number of DFFs acts as a parallel-input and parallel-output shift register. This structure of the shift register avoids changing the values of the control bits constantly during the interval when they are pushed-in serially. Eventually, the full-length *N*-bit shift register has an *N*-bit serial-in and parallel-out shift register in the bottom row and an *N*-bit parallel-in and parallel-out shift register in the top row.

The three input pins of the *N*-bit shift register have been named as BIT-IN, BIT-SHIFT, and BIT-LOAD. The serial bits from the microcontroller are provided to the shift register through the input pin, BIT-IN. Each serial bit is pushed into the shift register by providing a clock signal to the input pin, BIT-SHIFT, from the microcontroller. Once all the required number of serial bits have been pushed into the shift register (total *N* bits for the full-length shift register), all the bits are loaded simultaneously to the top row of the DFFs. The loading is done by providing a rising edge from the microcontroller through the pin BIT-LOAD. The outputs of the top row of the DFFs are then used as the control bits for the on-chip circuit.

References

- J. Proesel, C. Schow and A. Rylyakov, "25 Gb/s 3.6 pJ/b and 15 Gb/s 1.37 pJ/b VCSELbased optical links in 90 nm CMOS," *IEEE Int. Solid State Circuits Conference (ISSCC) Dig. Tech. Papers*, pp. 418-420, Feb. 2012.
- [2] D. Abts, M. R. Marty, P. M. Wells, P. Klausler and H. Liu, "Energy proportional datacenter networks," ACM 37th International Symposium on Computer Architecture (ISCA), Jun. 2010, pp. 338-347.
- [3] A. Roy, H. Zeng, J. Bagga, G. Porter and A. C. Snoeren, "Inside the social network's (Datacenter) network," ACM SIGCOMM Computer Communication Review, pp. 123-137, Aug. 2015.
- [4] M. Dayarathna, Y. Wen and R. Fan, "Data Center Energy Consumption Modeling: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 732-794, Firstquarter 2016.
- [5] M. Mansuri, J. E. Jaussi, J. E. Kennedy, T.-C. Hsueh, S. Shekhar, G. Balamurugan, F. O'Mahony, C. Roberts, R. Mooney and B. Casper, "A scalable 0.128–1 Tb/s, 0.8–2.6 pJ/bit, 64-lane parallel I/O in 32-nm CMOS," *IEEE J. of Solid-State Circuits*, vol. 48, no. 12, pp. 3229-3242, Dec. 2013.
- [6] B. Razavi, Design of Integrated Circuits for Optical Communication, New York: McGraw-Hill, 2003.
- [7] T. Anand, A. Elshazly, M. Talegaonkar, B. Young and P. K. Hanumolu, "A 5 Gb/s, 10 ns power-on-time, 36 μW off-state power, fast power-on transmitter for energy proportional links," *IEEE J. of Solid-State Circuits*, vol. 49, no. 10, pp. 2243-2258, Oct. 2014.
- [8] S. Sidiropoulos and M. A. Horowitz, "A semidigital dual delay-locked loop," *IEEE J. Solid-State Circuits*, vol. 32, no. 11, pp. 1683-1692, Nov. 1997.
- [9] Y.-H. Song, R. Bai, K. Hu, H.-W. Yang, P. Y. Chiang and S. Palermo, "A 0.47–0.66 pJ/bit, 4.8–8 Gb/s I/O transceiver in 65 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 48, no. 5, p. 1276–1289, May 2013.
- [10] B. Leibowitz, R. Palmer, J. Poulton, Y. Frans, S. Li, J. Wilson, M. Bucher, A. M. Fuller, J. Eyles, M. Aleksic, T. Greer and N. M. Nguyen, "A 4.3 Gb/s mobile memory interface with power-efficient bandwidth scaling," *IEEE J. Solid-State Circuits*, vol. 45, no. 4, p. 889–898, Apr. 2010.
- [11] M. Mansuri, J. E. Jaussi, J. E. Kennedy, T.-C. Hsueh, S. Shekhar, G. Balamurugan, F. O'Mahony, C. Roberts, R. Mooney and B. Casper, "A scalable 0.128-to-1Tb/s 0.8-to-2.6pJ/b 64-lane parallel I/O in 32nm CMOS," *IEEE Int. Solid State Circuits Conference (ISSCC) Dig. Tech. Papers*, Feb. 2013, pp. 402-403.
- [12] G. Balamurugan, J. Kennedy, G. Banerjee, J. E. Jaussi, M. Mansuri, F. O'Mahony, B. Casper and R. Mooney, "A Scalable 5–15 Gbps, 14–75 mW low-power I/O transceiver in 65 nm CMOS," *IEEE J. of Solid-State Circuits*, vol. 43, no. 4, pp. 1010-1019, Apr. 2008.
- [13] M. Hossain, K. Kaviani, B. Daly, M. Shirasgaonkar, W. Dettloff, T. Stone, K. Prabhu, B. Tsang, J. Eble and J. Zerbe, "A 6.4/3.2/1.6 Gb/s low power interface with all digital clock multiplier for on-the-fly rate switching," *Proceedings of the IEEE 2012 Custom Integrated Circuits Conference (CICC), San Jose, CA,* Sep. 2012, pp. 1-4.
- [14] C. Williams, G. E. R. Cowan and O. Liboiron-Ladouceur, "Power and noise configurable phase-locked loop using multi-oscillator feedback alignment," *IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug. 2013, pp. 1023-1026.
- [15] L. Rodoni, G. V. Buren, A. Huber, M. Schmatz and H. Jackel, "A 5.75 to 44 Gb/s Quarter Rate CDR With Data Rate Selection in 90 nm Bulk CMOS," *IEEE J. of Solid-State Circuits*, vol. 44, no. 7, pp. 1927-1941, Jul. 2009.
- [16] P. P. Dash, G. Cowan and O. Liboiron-Ladouceur, "A variable-bandwidth, power-scalable optical receiver front-end in 65 nm," in *IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug. 2013, pp. 717-720.
- [17] H.-S. Ko, L. M.-S. Sub-Han and S.-H. Chai, "1.25Gb/s burst-mode optical receiver for the ethernet PON using 0.35um CMOS technology," *The 6th International Conference on Advanced Communication Technology, Phoenix Park, Korea,* Feb. 2004, pp. 877-880.
- [18] T. Anand, M. Talegaonkar, A. Elkholy, S. Saxena, A. Elshazly and P. K. Hanumolu, "A 7 Gb/s embedded clock transceiver for energy proportional links," *IEEE J. of Solid-State Circuits*, vol. 50, no. 12, pp. 3101-3119, Dec. 2015.

- [19] T. Morf, M. Seifried, A. Cevrero, I. Ozkaya, C. Menolfi, D. Kuchta, M. Kossel, P. Francese, L. Kull, J. Kropp and T. Toifl, "VCSEL-based optical links in burst-mode slow optical power ramp-up and how to achieve ultra-short wake-up times," *Electronics Letters*, vol. 53, no. 19, pp. 1325-1327, Sep. 2017.
- [20] W. S. Choi, T. Anand, G. Shu, A. Elshazly and P. K. Hanumolu, ""A burst-mode digital receiver with programmable input jitter filtering for energy proportional links," *IEEE J. of Solid-State Circuits*, vol. 50, no. 3, pp. 737-748, Mar. 2015.
- [21] J. Lee and M. Liu, "A 20-Gb/s burst-mode clock and data recovery circuit using injectionlocking technique," *IEEE J. Solid-State Circuits*, vol. 43, no. 3, p. 619–630, Mar. 2008.
- [22] D. Dunwell, A. Carusone, J. Zerbe, B. Leibowitz, B. Daly and J. Eble, "A 2.3–4GHz injection-locked clock multiplier with 55.7% lock range and 10-ns power-on," *IEEE Custom Integrated Circuits Conf. (CICC)*, Sep. 2012, pp. 1–4.
- [23] A. Rylyakov, J. Proesel, S. Rylov, B. Lee, J. Bulzacchelli, A. Ardey, B. Parker, M. Beakes,
 C. Baks, C. Schow and M. Meghelli, "A 25 Gb/s burst-mode receiver for rapidly reconfigurable optical networks," *IEEE Int. Solid State Circuits Conference (ISSCC) Dig. Tech. Papers*, Feb. 2015, pp. 400-401.
- [24] T. Anand, M. Talegaonkar, A. Elshazly, B. Young and P. K. Hanumolu, "A 2.5 GHz 2.2mW/25 μW on/off-state power 2psrms-long-term-jitter digital clock multiplier with 3-reference-cycles power-on time," *IEEE Int. Solid State Circuits Conference (ISSCC) Dig. Tech. Papers*, Feb. 2013, pp. 256–257.
- [25] T. Toifl, C. Menolfi, P. Buchmann, C. Hagleitner, M. Kossel, T. Morf, J. Weiss and M. Schmatz, "A 72mW 0.03mm2 Inductorless 40Gb/s CDR in 65nm SOI CMOS," *IEEE Int. Solid State Circuits Conference (ISSCC) Dig. Tech. Papers*, Feb. 2007, pp. 226-227.
- [26] J. Lee and B. Razavi, "A 40-Gb/s Clock and Data Recovery Circuit in 0.18-μm CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2181-2190, Dec. 2003.
- [27] C. Kromer, G. Sialm, C. Menolfi, M. Schmatz, F. Ellinger and H. Jackel, "A 25-Gb/s CDR in 90-nm CMOS for High-Density Interconnects," *IEEE Int. Solid State Circuits Conference* (ISSCC) Dig. Tech. Papers, Feb. 2006, pp. 326-327.

- [28] I. Ozkaya, A. Cevrero, P. A. Francese, C. Menolfi, M. Braendli, T. Morf, D. Kuchta, L. Kull, M. Kossel, D. Luu, M. Meghelli, Y. Leblebici and T. Toifl, "A 56Gb/s burst-mode NRZ optical receiver with 6.8ns power-on and CDR-Lock time for adaptive optical links in 14nm FinFET CMOS," *IEEE Int. Solid State Circuits Conference (ISSCC) Dig. Tech. Papers*, Feb. 2018, pp. 266–268.
- [29] A. K. M. D. Hossain, Aurangozeb and M. Hossain, "Burst mode optical receiver with 10 ns lock time based on concurrent DC offset and timing recovery technique," *IEEE Journal of Optical Communications and Networking*, vol. 10, no. 2, pp. 65-78, Feb. 2018.
- [30] A. B. Mazwar, T. Kuriyama and H. Ueda, "10 Gbps optical burst mode receiver with fast response and high stability," *IEEE International Conference on Communications (ICC)*, May 2016, pp. 1-6.
- [31] A. Ragab, Y. Liu, K. Hu, P. Chiang and S. Palermo, "Receiver jitter tracking characteristics in high-speed source synchronous links," *J. Electrical and Computer Engineering*, vol. 2011, pp. 1-15, 2011.
- [32] R. Kreienkamp, U. Langmann, C. Zimmermann, T. Aoyama and H. Siedhoff, "A 10-Gb/s CMOS clock and data recovery circuit with an analog phase interpolator," *IEEE J. Solid-State Circuits*, vol. 40, no. 3, pp. 736-743, Mar. 2005.
- [33] J. Kim and J. F. Buckwalter, "Bandwidth enhancement with low group-delay variation for a 40-Gb/s transimpedance amplifier," *IEEE Transactions on Circuit and Systems I*, vol. 57, no. 8, pp. 1964-1972, Aug. 2010.
- [34] S. Palermo, "Design of High-Speed Optical Interconnect Transceiver," Ph. D. dissertation, Dept. Elect. Eng., Standford Univ., 2007.
- [35] J. D. H. Alexander, "Clock recovery from random binary data," *Electronics Letters*, vol. 11, pp. 541-542, Oct .1975.
- [36] K.-S. Park, B.-J. Yoo, M.-S. Hwang, H. Chi, H.-C. Kim, J.-W. Park, K. Kim and D.-K. Jeong,
 "A 10 Gb/s optical receiver front-end with 5-mW transimpedance amplifier," *IEEE Asian Solid State Circuits Conference*, Nov. 2010, pp. 1-4.

- [37] Y.-H. Chien, K.-L. Fu and S.-I. Liu, "A 3-25 Gb/s four-channel receiver with noise-canceling TIA and power-scalable LA," *IEEE Transactions on Circuits and Systems II*, vol. 61, no. 11, pp. 845-849, Nov. 2014.
- [38] L. Szilagyi, J. Pliva, R. Henker, D. Schoeniger, J. P. Turkiewicz and F. Ellinger, "A 53-Gbit/s optical receiver frontend with 0.65 pJ/bit in 28-nm bulk-CMOS," *IEEE J. Solid-State Circuits*, vol. 54, no. 3, pp. 845-855, Mar. 2019.
- [39] M. M. P. Fard, O. Liboiron-Ladouceur and G. Cowan, "1.23-pJ/bit 25-Gb/s inductor-less optical receiver with low-voltage silicon photodetector," *IEEE J. Solid-State Circuits*, vol. 53, no. 6, pp. 1793-1805, Jun. 2018.
- [40] A. Agrawal, A. Liu, P. K. Hanumolu and G. Y. Wei, "An 8 x 5 Gb/s parallel receiver with collaborative timing recovery," *IEEE J. Solid-State Circuits*, vol. 44, no. 11, pp. 3120-3130, Nov. 2009.
- [41] A. Cevrero, I. Ozkaya, P. A. Francese, C. Menolfi, M. Braendli, T. Morf, D. Kuchta, M. Kossel, L. Kull, D. Luu, J. Proesel, Y. Leblebici and T. Toifl, "A 60 Gb/s 1.9 pJ/bit NRZ optical-receiver with low latency digital CDR in 14nm CMOS FinFET," *Symposium on VLSI Circuits*, Jun. 2017, pp. 320-321.
- [42] H. Wang and A. Hajimiri, "A wideband CMOS linear digital phase rotator," *IEEE Custom Integrated Circuits Conf. (CICC)*, Sep. 2017, pp. 671-674.
- [43] P. K. Hanumolu, G.-Y. Wei and U.-K. Moon, "A wide-tracking range clock and data recovery circuit," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 425-439, Feb. 2008.
- [44] D. Kim, M. G. Ahmed, W.-S. Choi, A. Elkholy and P. K. Hanumolu, "A 12-Gb/s 10-ns Turn-On Time Rapid ON/OFF Baud-Rate DFE Receiver in 65-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 55, no. 8, p. 2196–2205, Aug. 2020.
- [45] E. Säckinger, Broadband Circuits for Optical Fiber Communication, Hoboken, NJ, USA: Wiley, 2005.
- [46] M. Banu and A. E. Dunlop, "Clock recovery circuits with instantaneous locking," *Electronics Letters*, vol. 28, no. 23, pp. 2127-2130, Nov. 1992.