A 1.8-pJ/b 16x16-Gb/s Source-Synchronous Parallel Interface in 32nm SOI CMOS with Receiver Redundancy for Link Recalibration

Timothy O. Dickson, Yong Liu*, Ankur Agrawal, John F. Bulzacchelli, Herschel Ainspan, Zeynep Toprak-Deniz, Benjamin D. Parker, Mounir Meghelli, and Daniel J. Friedman

IBM Research Division
T. J. Watson Research Center
Yorktown Heights, NY 10598 USA

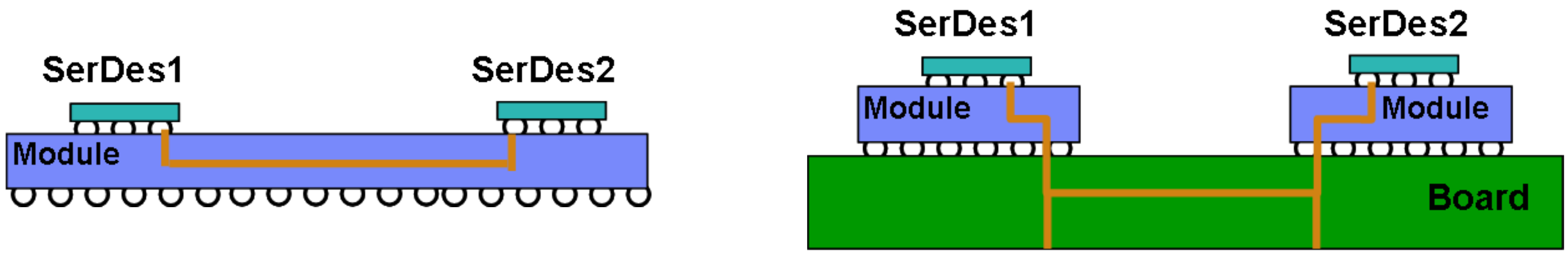
*** Now with Jiao Tong University, Shanghai, China**

Outline

- Introduction & Background
- Receiver Redundancy Architecture
- Transceiver Test Chip Design
- Measurement Results
- Conclusions

Introduction

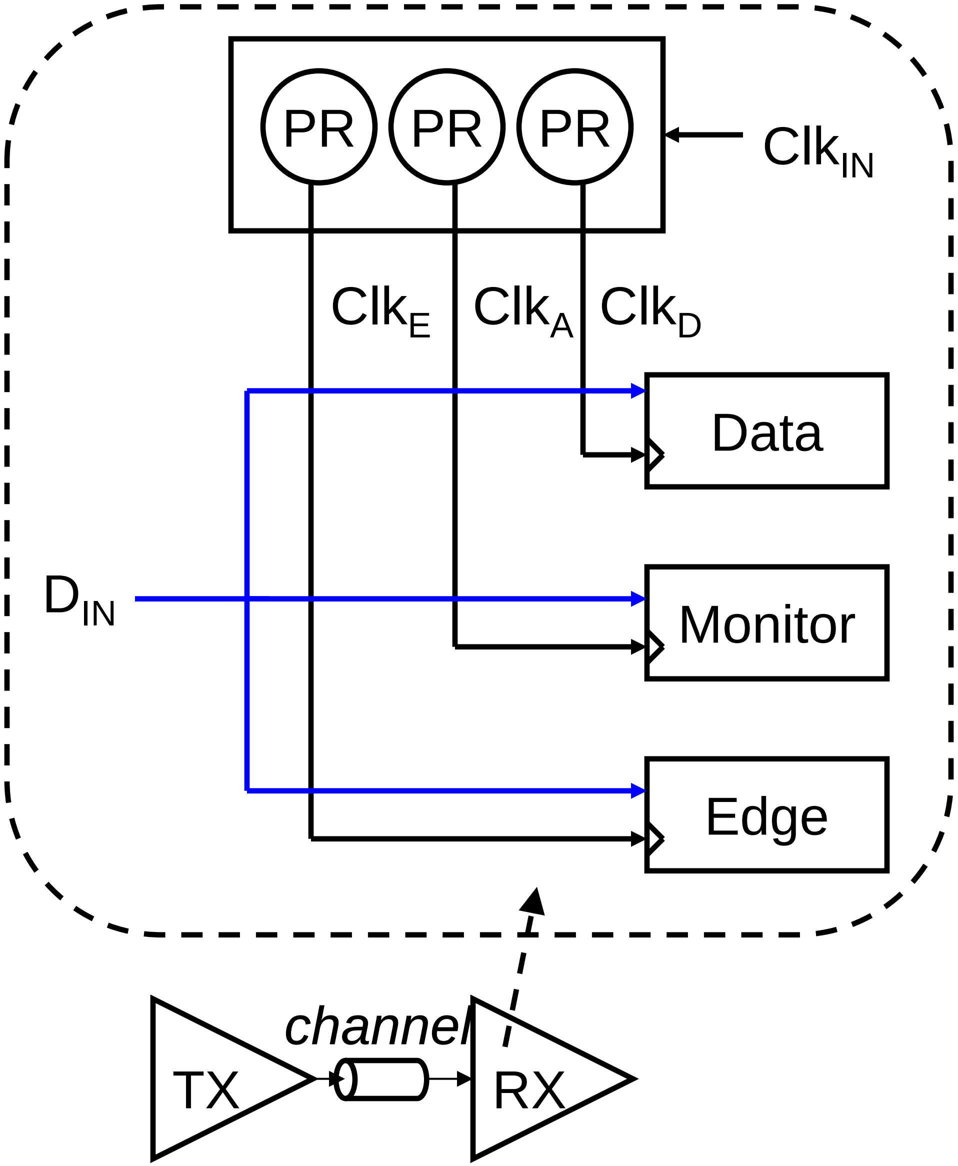
- **Application:** 10-15Gb/s Ultra-short reach (USR) source synchronous I/O for processor-to-processor or processor-to-memory links
  - Chip-to-chip interconnects on an MCM
  - Chip-to-chip links on a low-loss (<10dB) PCB



- **Example:** Hybrid Memory Cube 10G USR interface
  - 320GB/s aggregate link bandwidth
  - Highly efficient I/O (<2pJ/bit) are required to support high-density, high-bandwidth memory
- **Challenge:** Making I/O robust to changes in operating conditions and enabling periodic recalibration with minimal power/area overhead

Architecture innovations are required to support link calibration with minimal circuitry overhead.

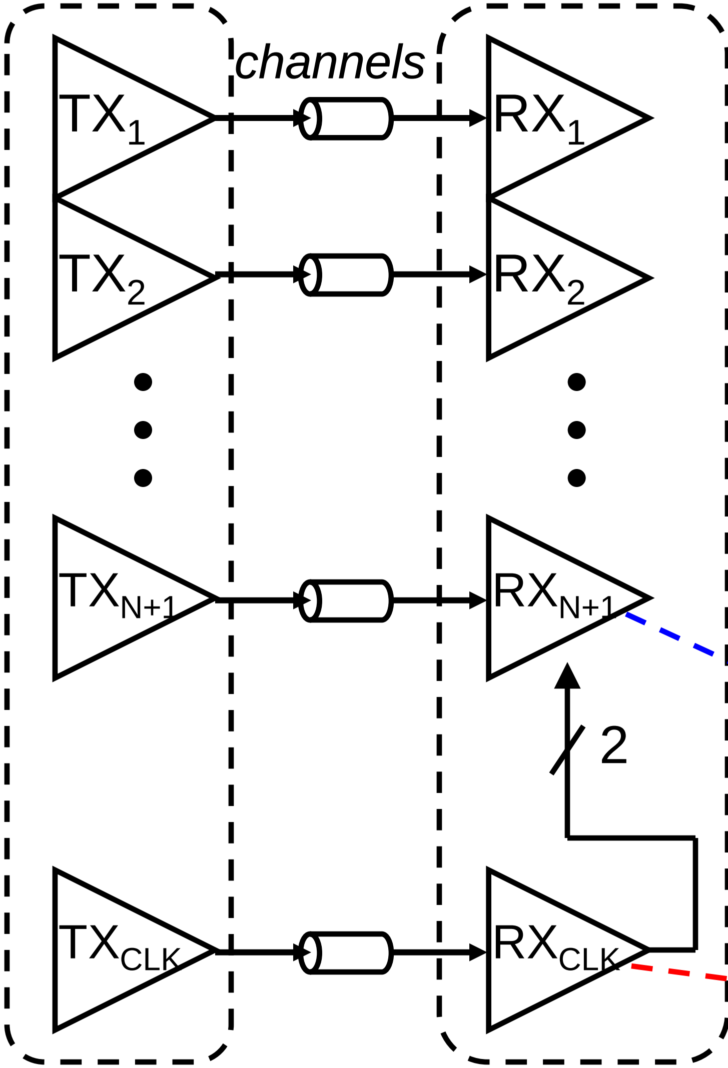
Conventional Serial Receiver with Equalizer Adaptation



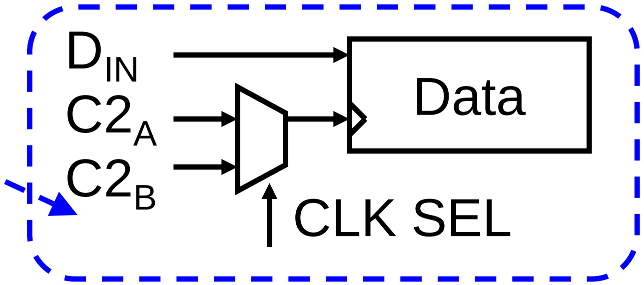
- Three sets of samplers
  - Data
  - Edge for CDR
  - Eye monitor needed to generate error signals used by equalizer adaptation algorithm (e.g., SS-LMS)
- Phase rotators needed to generate clock phases for data, edge, and monitor samplers.

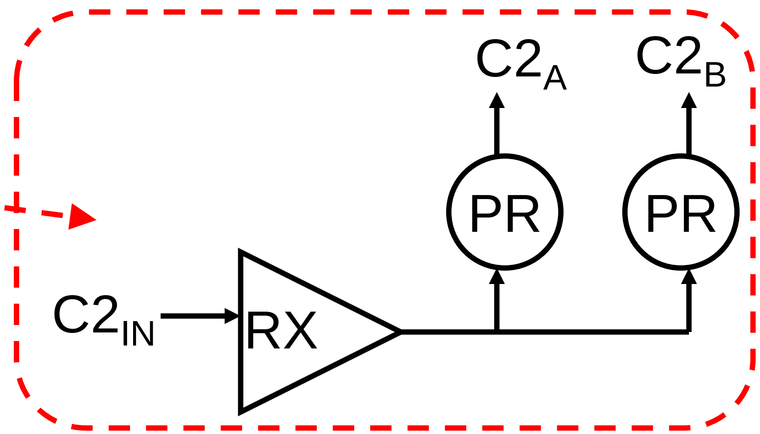
200% circuitry overhead is prohibitive for high-density I/O applications

<u>Prior Art:</u> Link Redundancy for High-Density Interconnects



- Move redundancy up in the bus hierarchy
  - Use N+1 lanes to transmit N bits of data
  - Extra lane enables round-robin link calibration
  - Known pattern (e.g. PRBS) must be sent across the link being calibrated
  - Dual phase rotator architecture

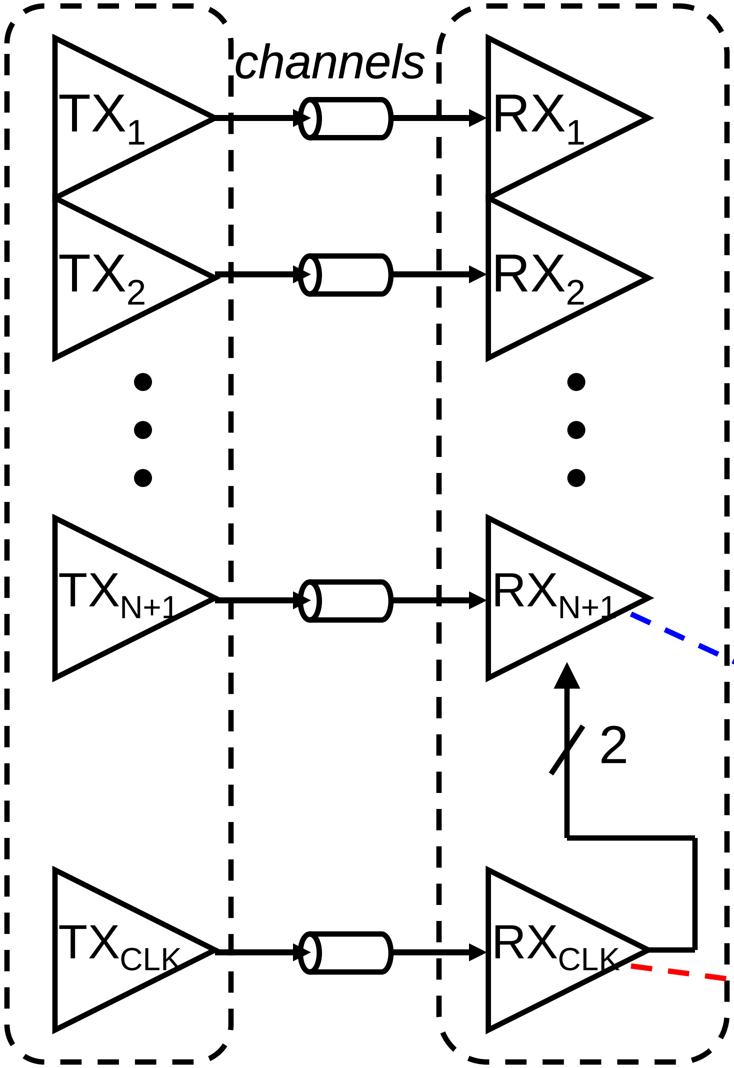




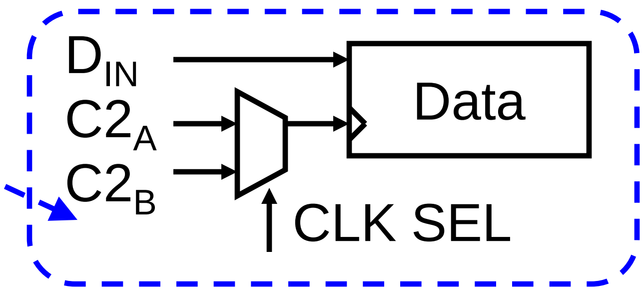
Second phase rotator enables eye scan during calibration

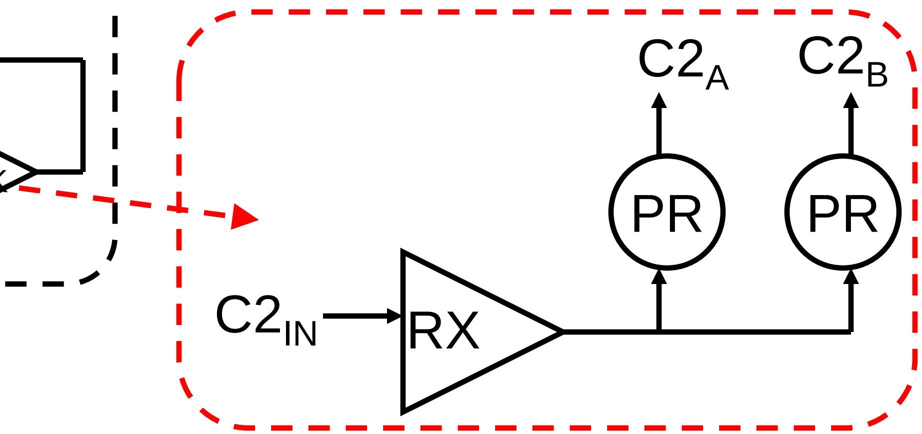
Dickson et al., VLSI 2011 & JSSC April 2012

Prior Art: Link Redundancy for High-Density Interconnects



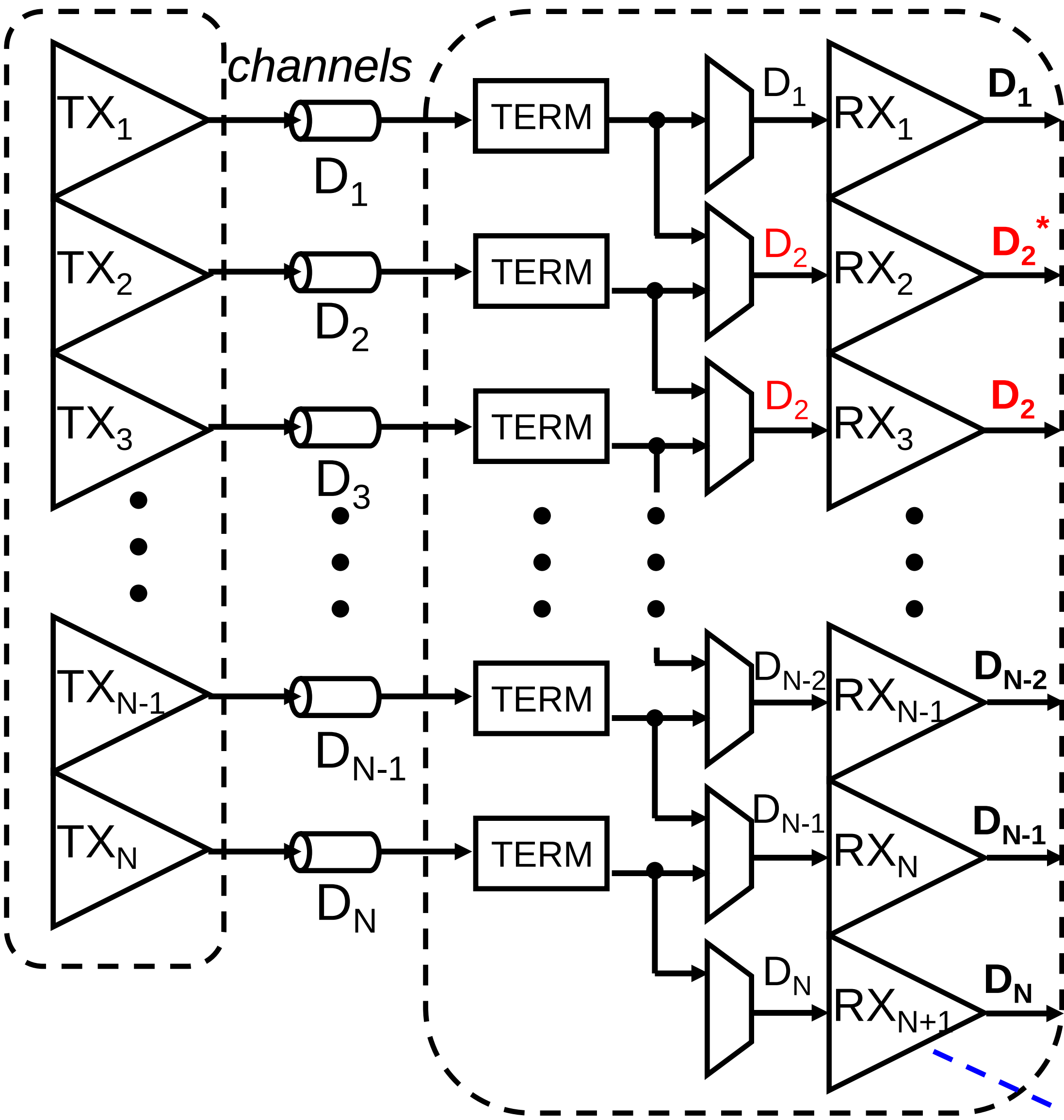
- **Disadvantages**
  - Additional lanes (channels, C4s, package traces, etc) to support calibration
  - Scheduling between transmitter and receiver
  - Power overhead associated with extra transmitter



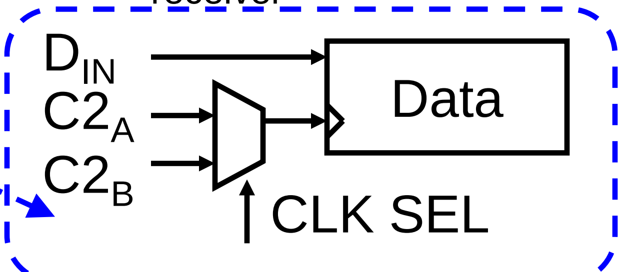


Dickson et al., VLSI 2011 & JSSC April 2012

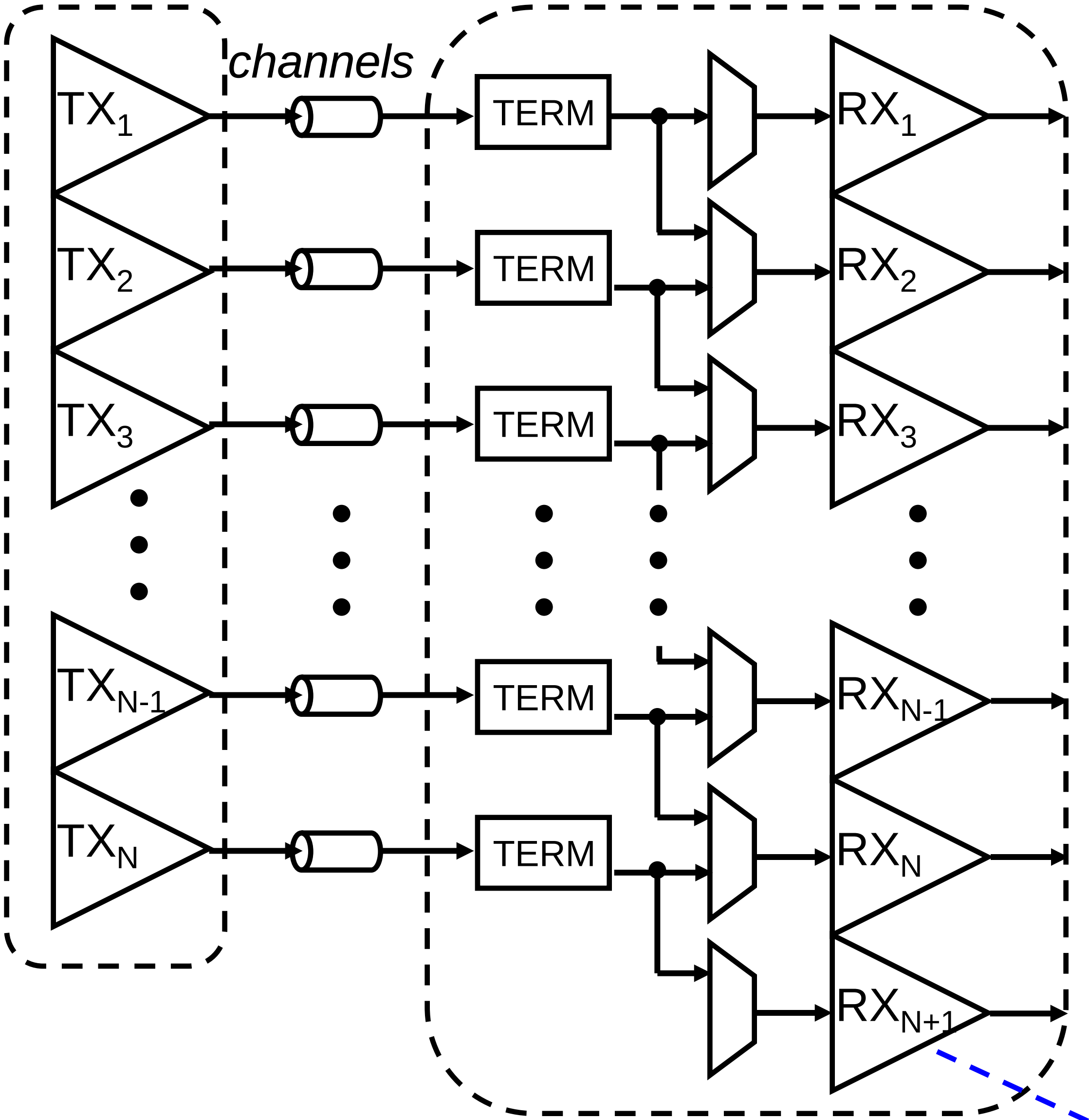
Receiver Redundancy: Concept



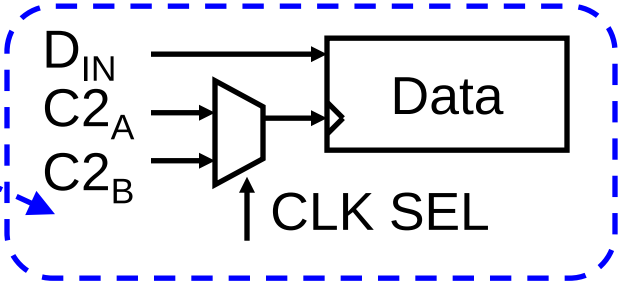
- Transmit N bits of data using N transmitters and N+1 receivers
- Calibrate receivers one at a time in a round-robin fashion, with the other N receivers processing data uninterrupted
- Data assumptions can be qualified using outputs from the adjacent receiver
  - Example – Calibrating RX_2
- As with link redundancy, two clocks must be distributed across the bus (not shown)
  - One 'fixed' for sampling data in N receivers
  - One 'adjustable' for calibration of 1 receiver



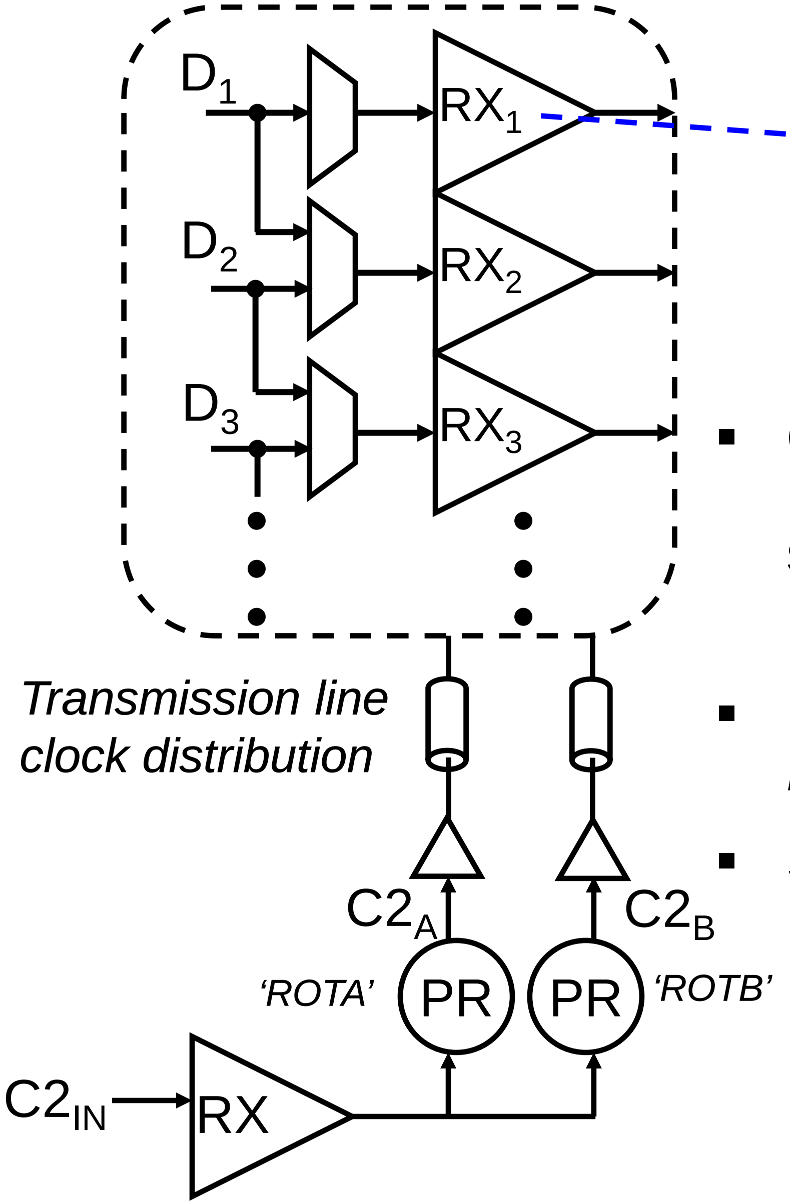
Receiver Redundancy: Considerations

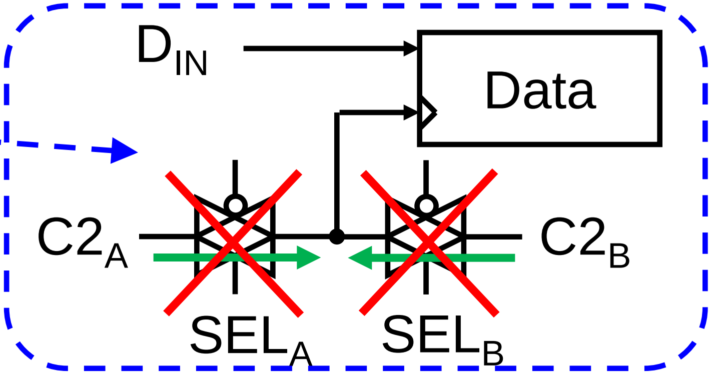


- Handoff of data between receivers must be seamless and error-free
- Handoff between clocks within a receiver must be glitchless



Glitchless Clock Handoff



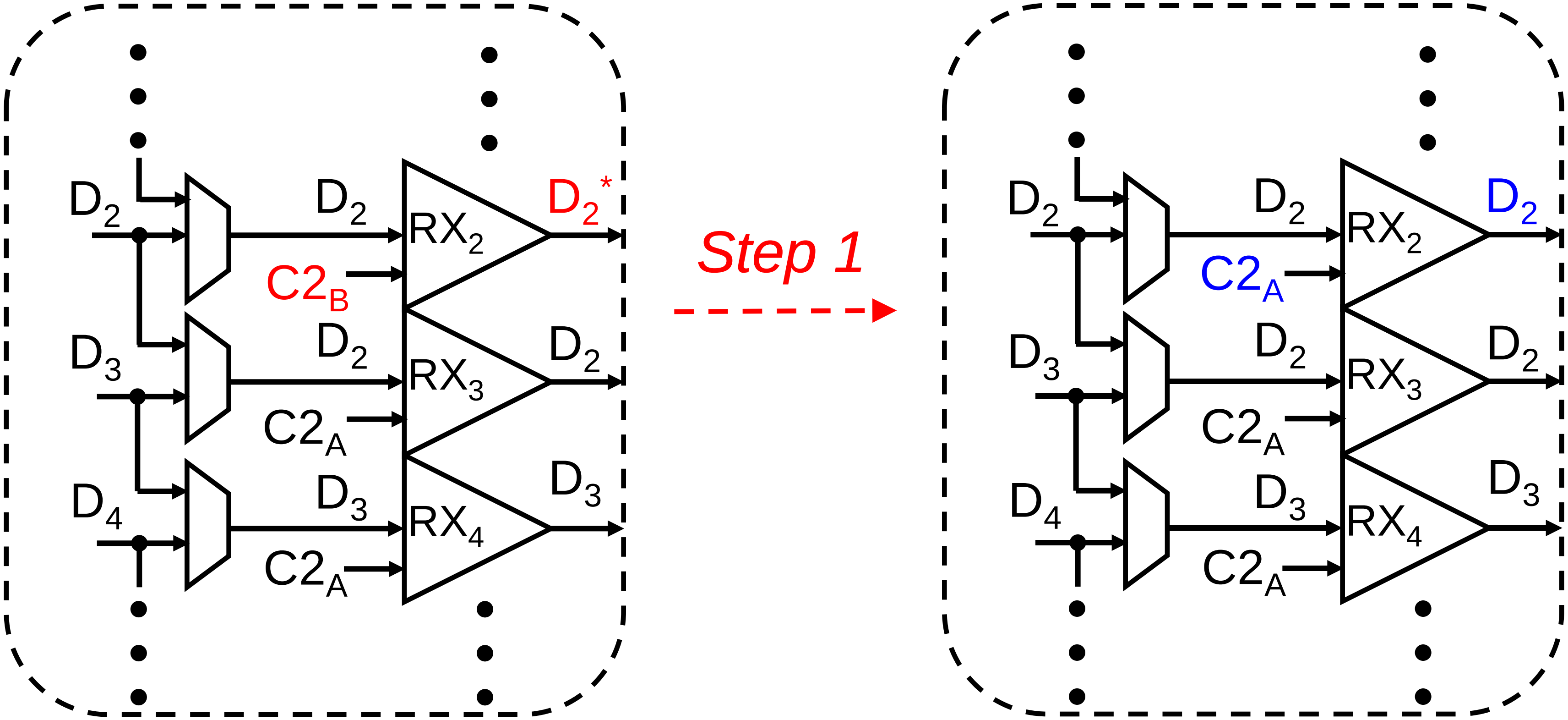


- Clocks are distributed to all receivers using a low-skew (<2ps) transmission line distribution scheme
  - Dickson et al. CICC 2014 & JSSC Aug 2015
- Phase offset between two phase rotators *ROTA* and *ROTB* are determined at startup
- Switching RX glitchlessly from $C2_A$ to $C2_B$:
  1. Adjust phase rotator *ROTB* such that $C2_B$ is in phase with $C2_A$
  2. Within the receiver, close both transmission gates in the clock select
  3. Within the receiver, open the transmission gate associated with $C2_A$

Data Handoff Example

Scenario: Return RX_2 into service, take RX_3 out of service

Example assumption: $C2_A$ represents the fixed sampling clock
$C2_B$ represents the adjustable clock for calibration

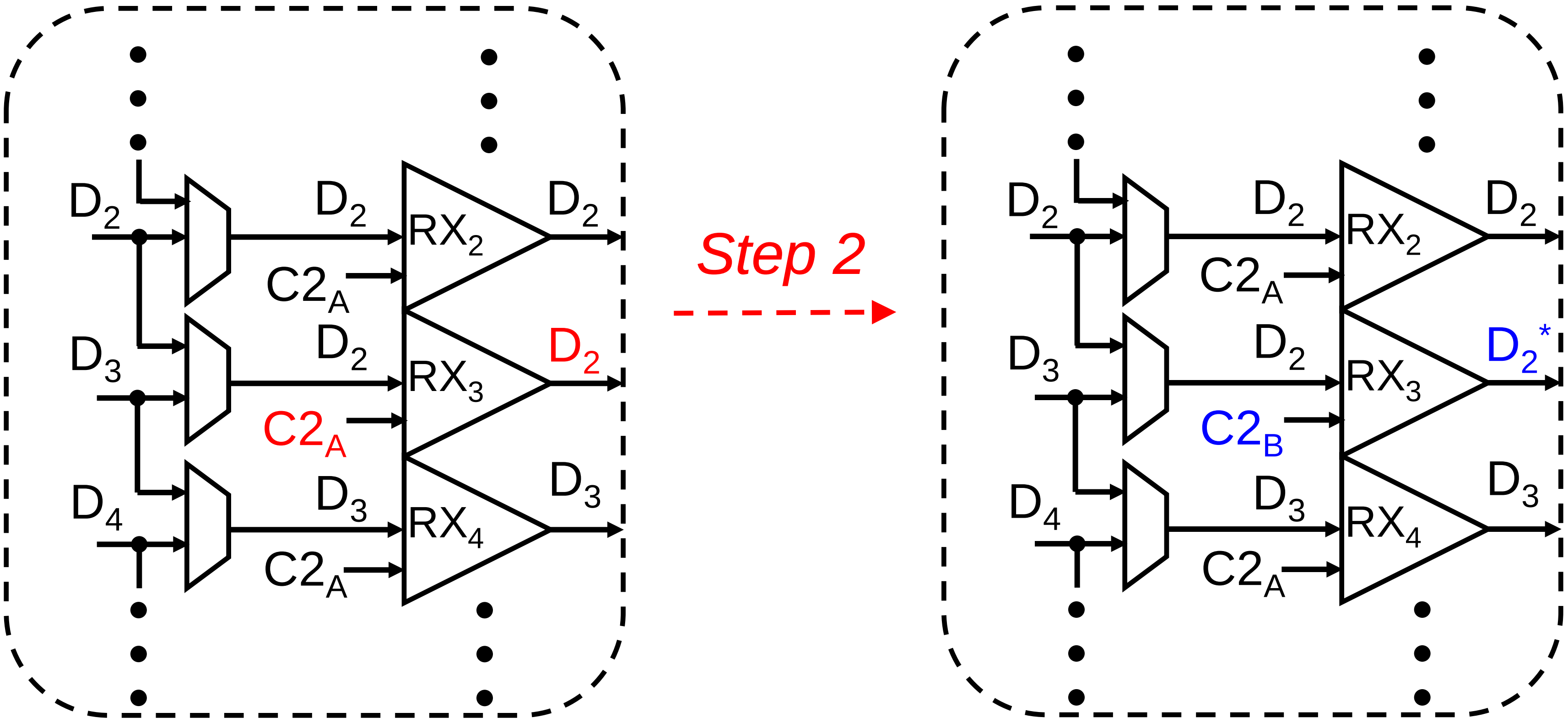


Switch clock input on RX_2 from $C2_B$ to $C2_A$
RX_2 can then be placed back in service, receiving data D_2

Data Handoff Example

Scenario: Return RX_2 into service, take RX_3 out of service

Example assumption: $C2_A$ represents the fixed sampling clock
$C2_B$ represents the adjustable clock for calibration

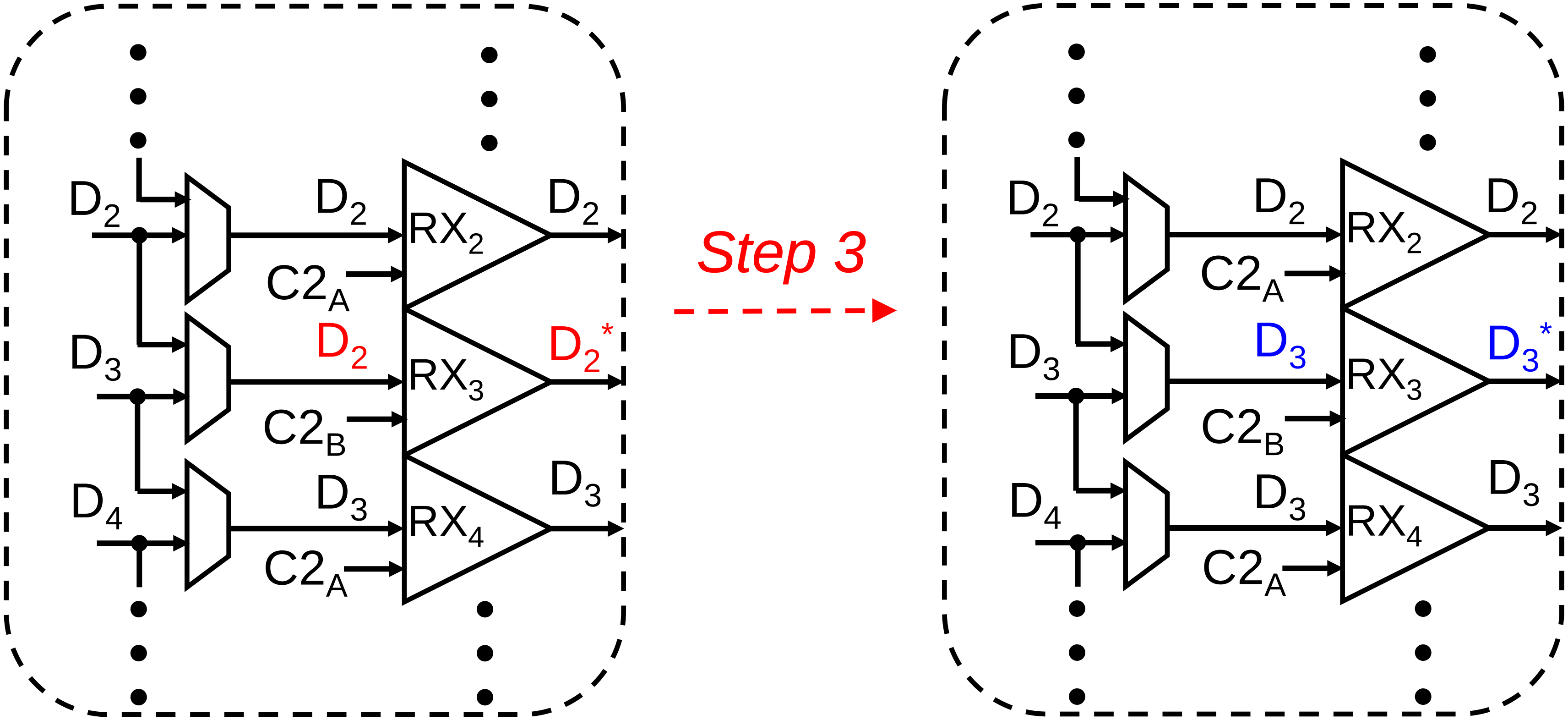


Switch clock input on RX_3 from $C2_A$ to $C2_B$

Data Handoff Example

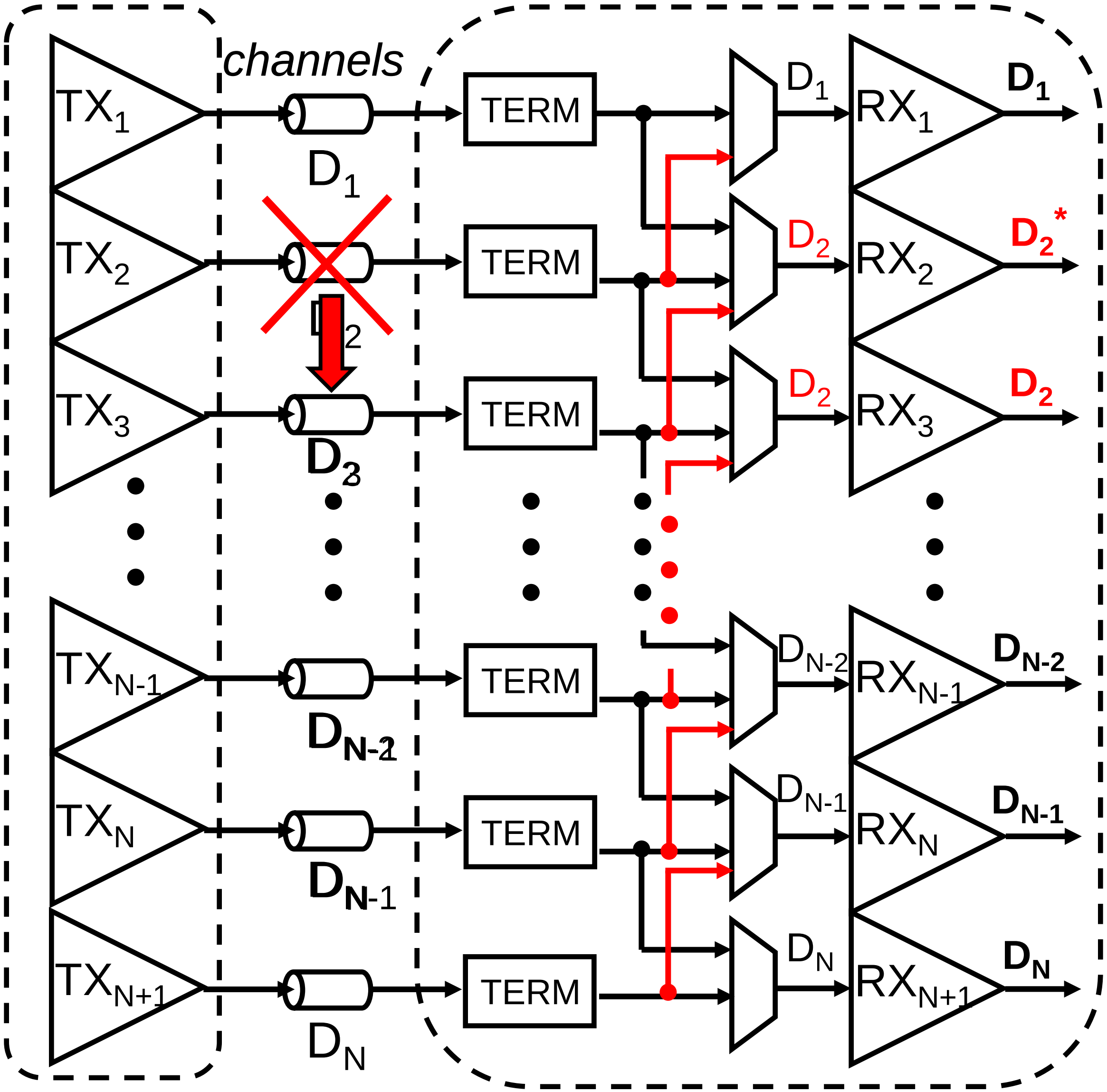
Scenario: Return RX_2 into service, take RX_3 out of service

Example assumption: $C2_A$ represents the fixed sampling clock
$C2_B$ represents the adjustable clock for calibration



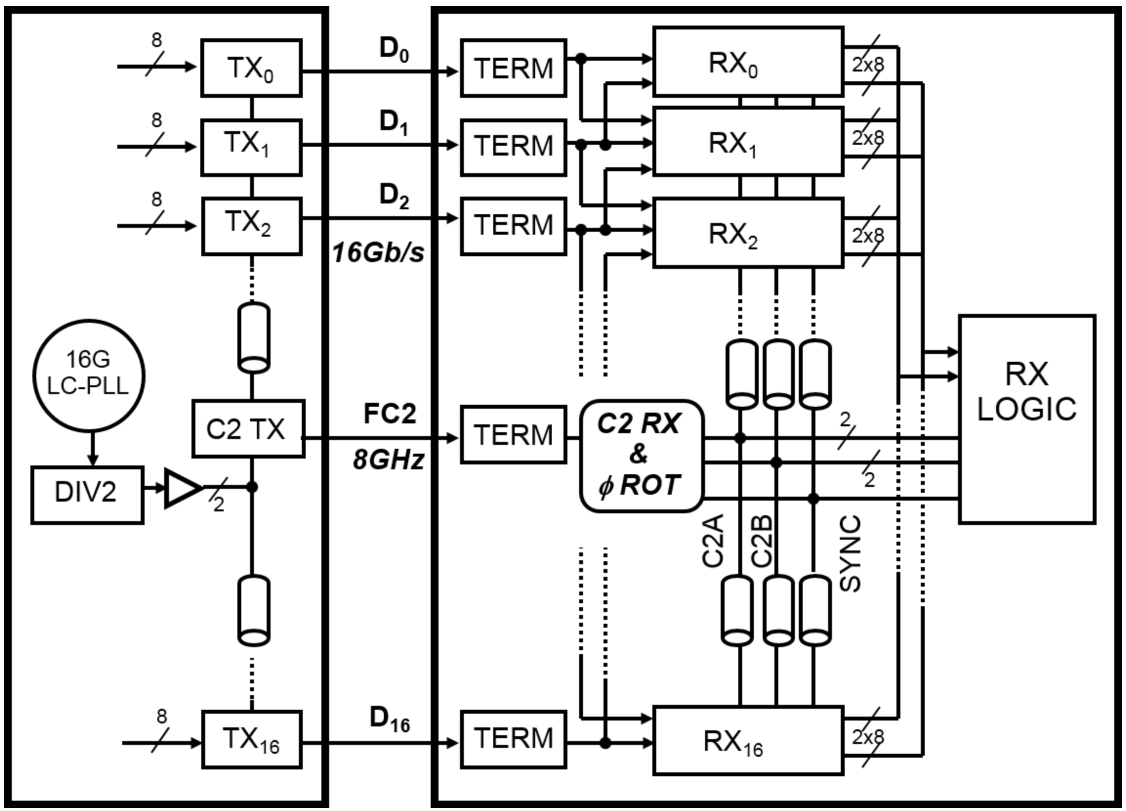
Switch data input on RX_3 from D_2 to D_3

Enabling Data Spares



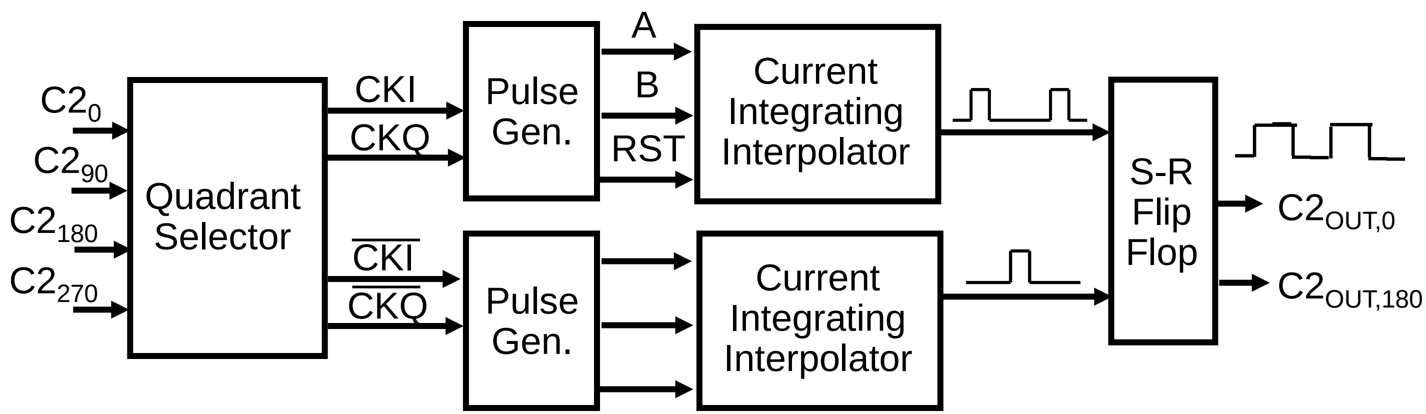
- Spare lane still required to facilitate communication of N bits of data in the presence of mechanical faults (e.g., open circuits, out-of-spec channels)
- In illustrated example, data D_2 cannot be coupled to receiver RX_2.
  - Third input to RX multiplexer is required to support calibration in the presence of faults

16 x 16Gb/s Test Chip in 32nm SOI CMOS

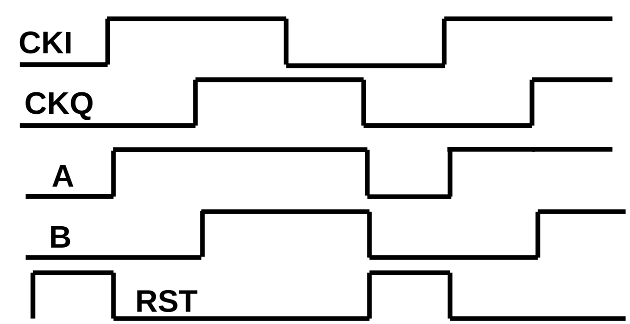


- Integrated transceiver test chip designed in 32nm SOI CMOS
- 16 data lanes plus 1 spare
  - Demultiplexed RX data distributed to shared RX logic
- Half-rate (8 GHz) forwarded clock FC2
  - $C2_A$ & $C2_B$ distributed to RXs, along with SYNC pulse

Phase Rotator Architecture

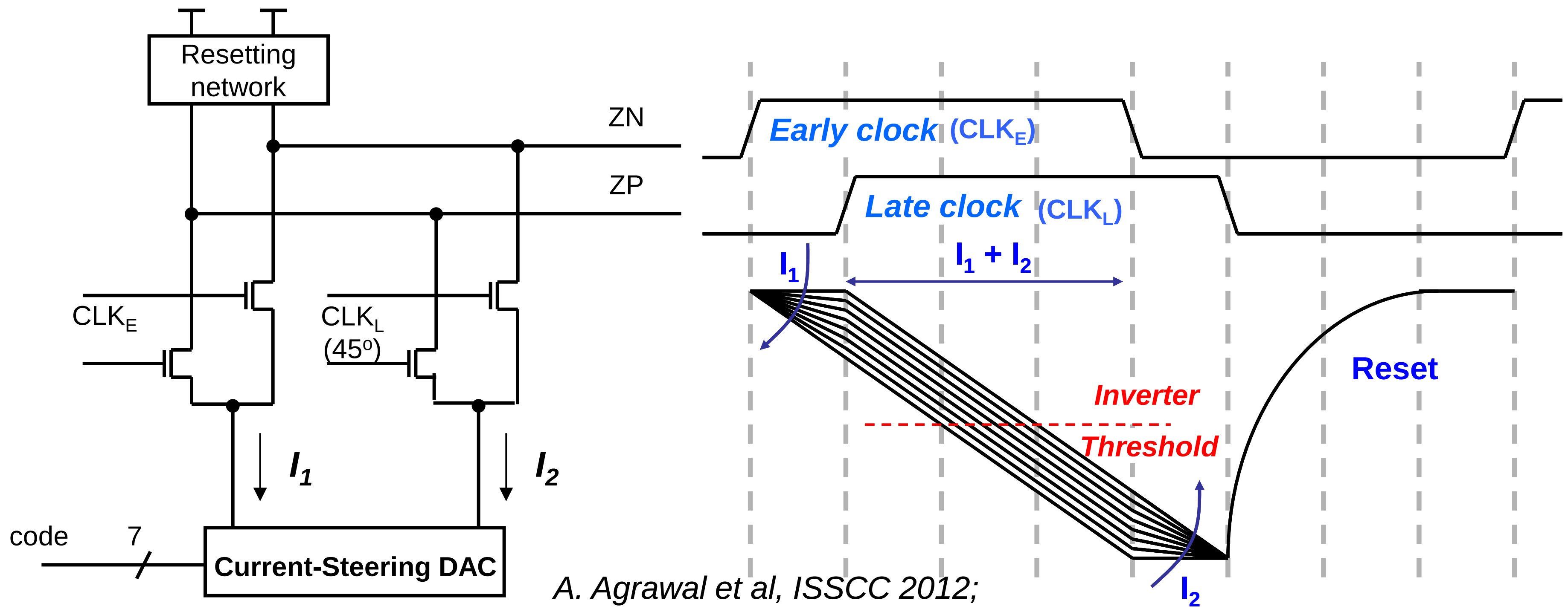


- 8-GHz, 6b phase rotator operates on CMOS rail-to-rail I/Q clocks
- Pulse generator circuit creates early (A) and late (B) trigger signal for 4b current integrating phase interpolators
- SR flip-flop cleans up clock duty cycle after integrating interpolators



Pulse Generator Timing

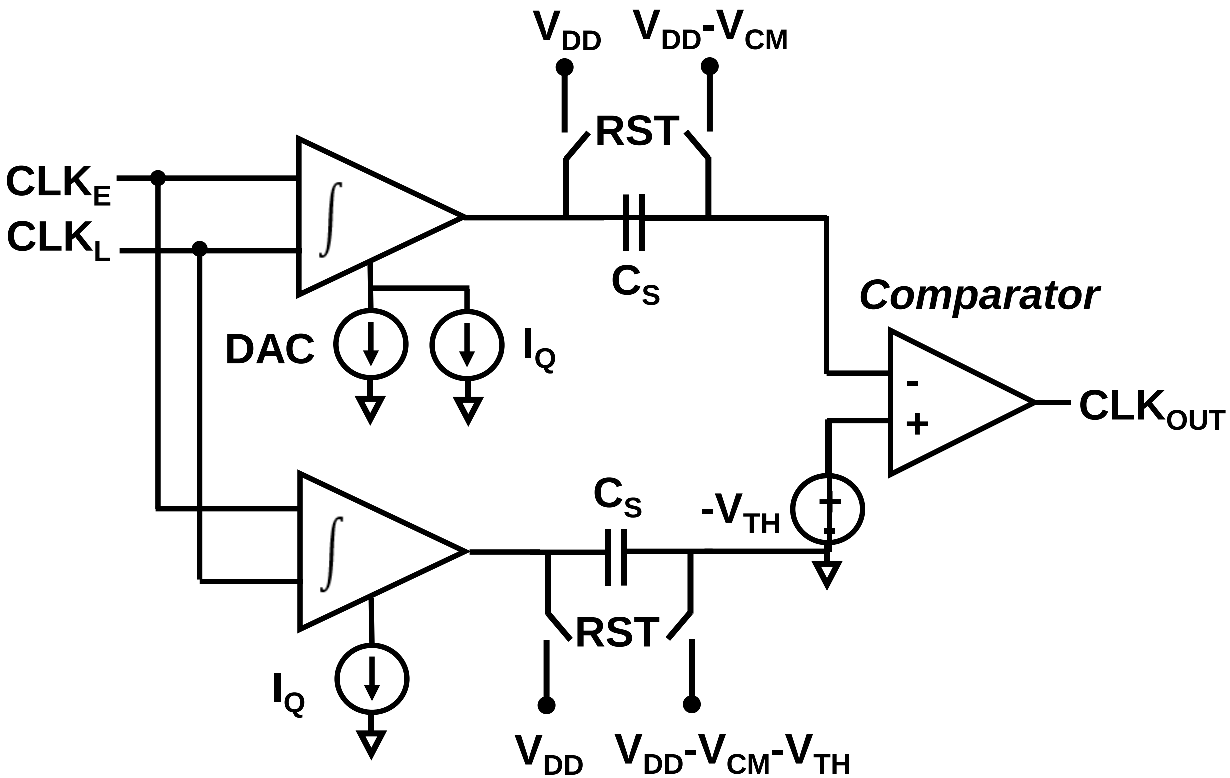
Prior Art: Current-Integrating Phase Interpolator



A. Agrawal et al, ISSCC 2012;
Dickson et al, CICC 2014 & JSSC Aug 2015

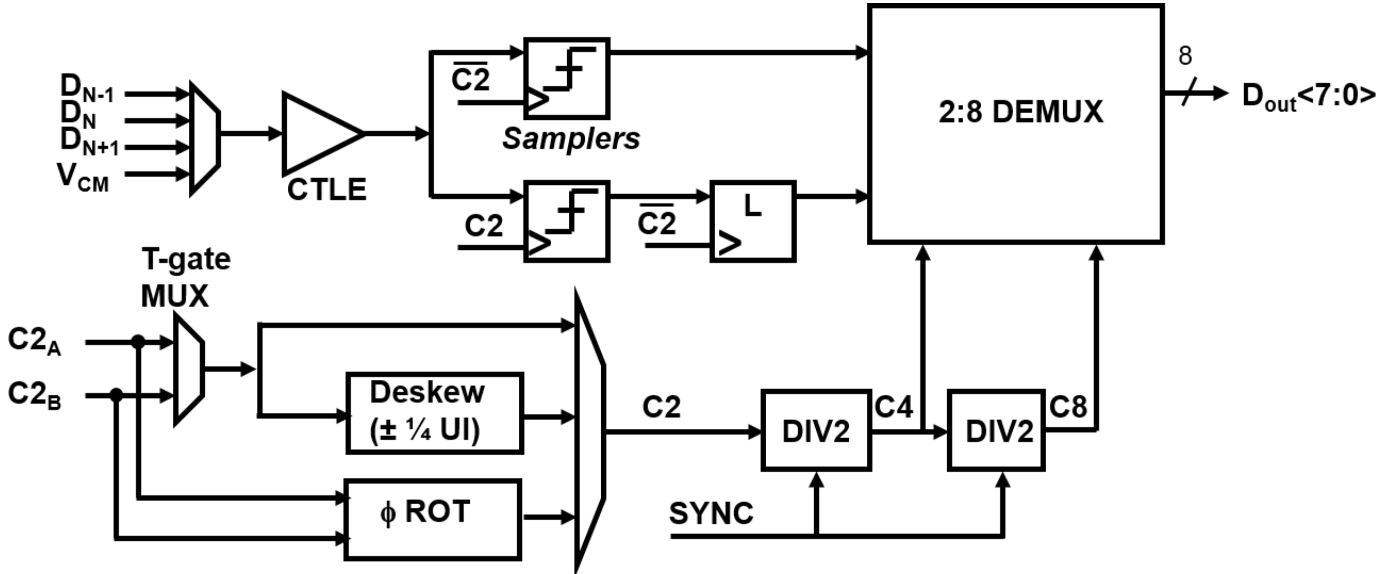
- Early clock triggers integration of variable current I_1 from current-steering DAC onto output capacitance
- Late clock triggers remainder I_2 of DAC current to add to I_1
  - Time at which loading inverter (not shown) is triggered depends on rotator code
- Disadvantage:
  - Outputs drive psuedo-differential CMOS inverters, making the approach susceptible to power supply noise

Modified Current-Integrating Phase Interpolator



- <u>Key Architecture Improvement:</u> Differential comparator detects when integrator output crosses threshold
- Replica integrator operates on quiescent current
- Capacitive level shifter introduces threshold and sets proper input common-mode level for comparator

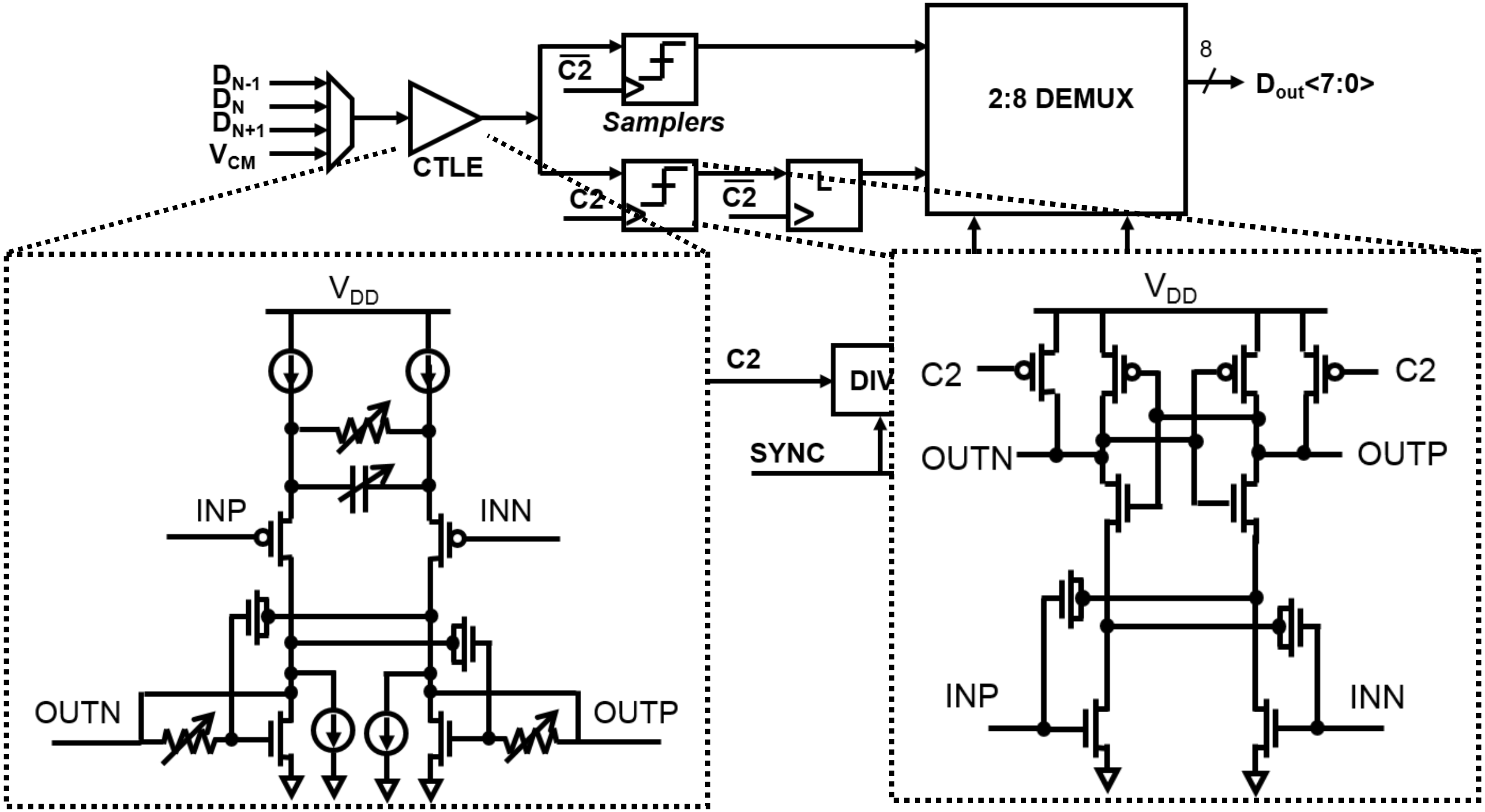
Data Receiver Lane



- Data path:
  - CTLE with 9dB of peaking at 8GHz
  - Analog MUX to support receiver sparing architecture
  - Low latency 2:8 demultiplexer
- Clock path:
  - Three deskew modes for variety of channel applications*
  - C4 and C8 dividers flushed using global SYNC signal to synchronize dividers in all lanes

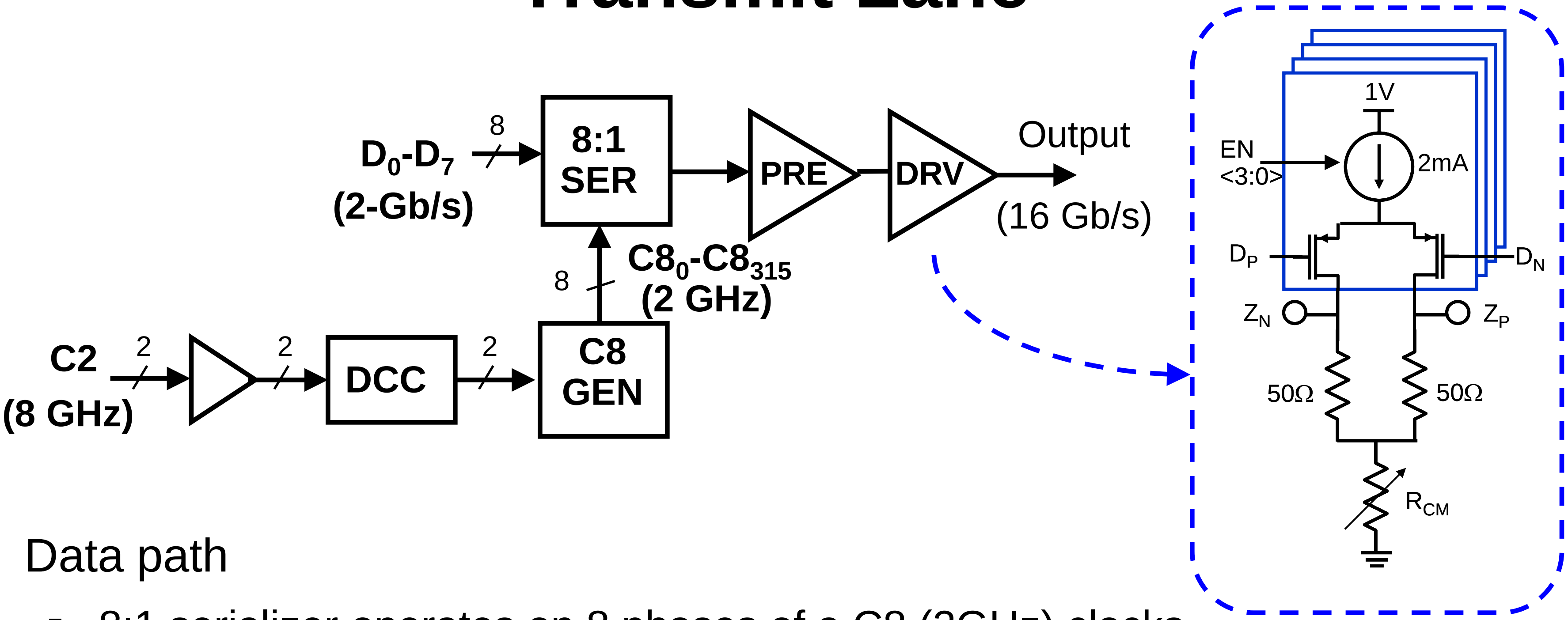
** Dickson et al, CICC 2014 & JSSC Aug 2015*

CTLE and Samplers



- CTLE:
  - Active inductor to boost high-frequency gain
  - Parallel current sources in the load improve G_m of the input pair relative to the active load
- Sampler:
  - Input pair forms a current mirror with the CTLE active inductor load

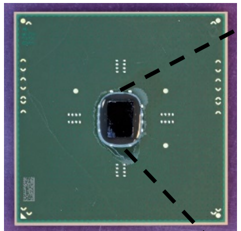
Transmit Lane

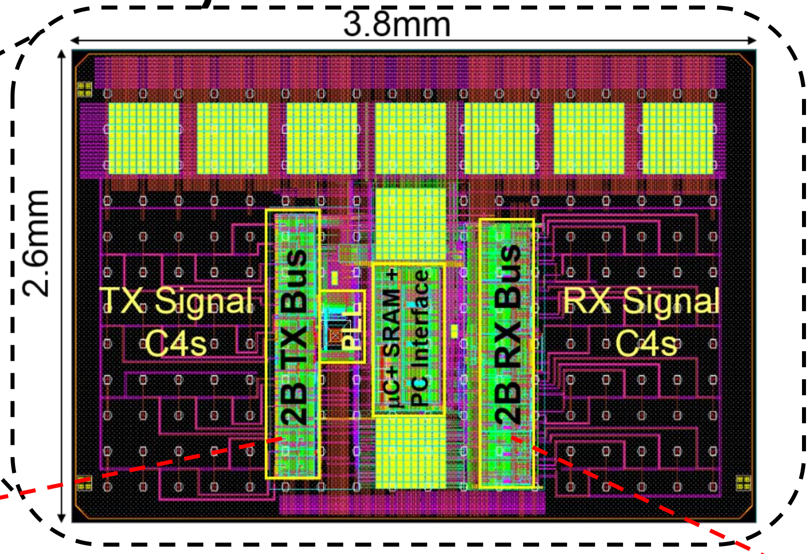


- Data path
  - 8:1 serializer operates on 8 phases of a C8 (2GHz) clocks
  - CML driver with 2b amplitude control (100-400mVppd)
- Clock path
  - C2 clock received from transmission-line clock distribution
  - Cascade of I/Q dividers produces C8 clock phases for serializer

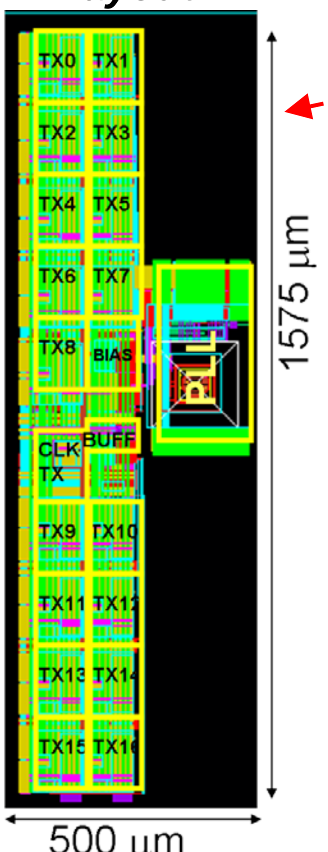
Layout, Fabrication, and Module Assembly

24 x 24mm² SLC

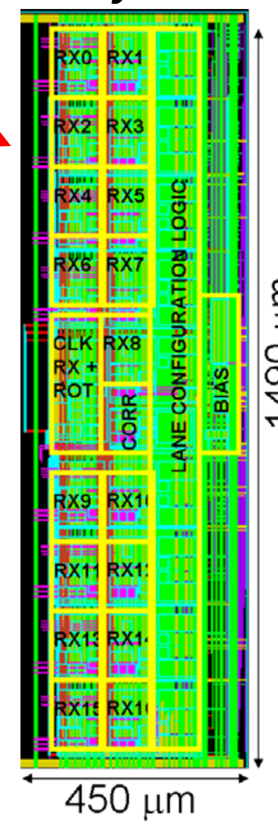




TX Bus Layout

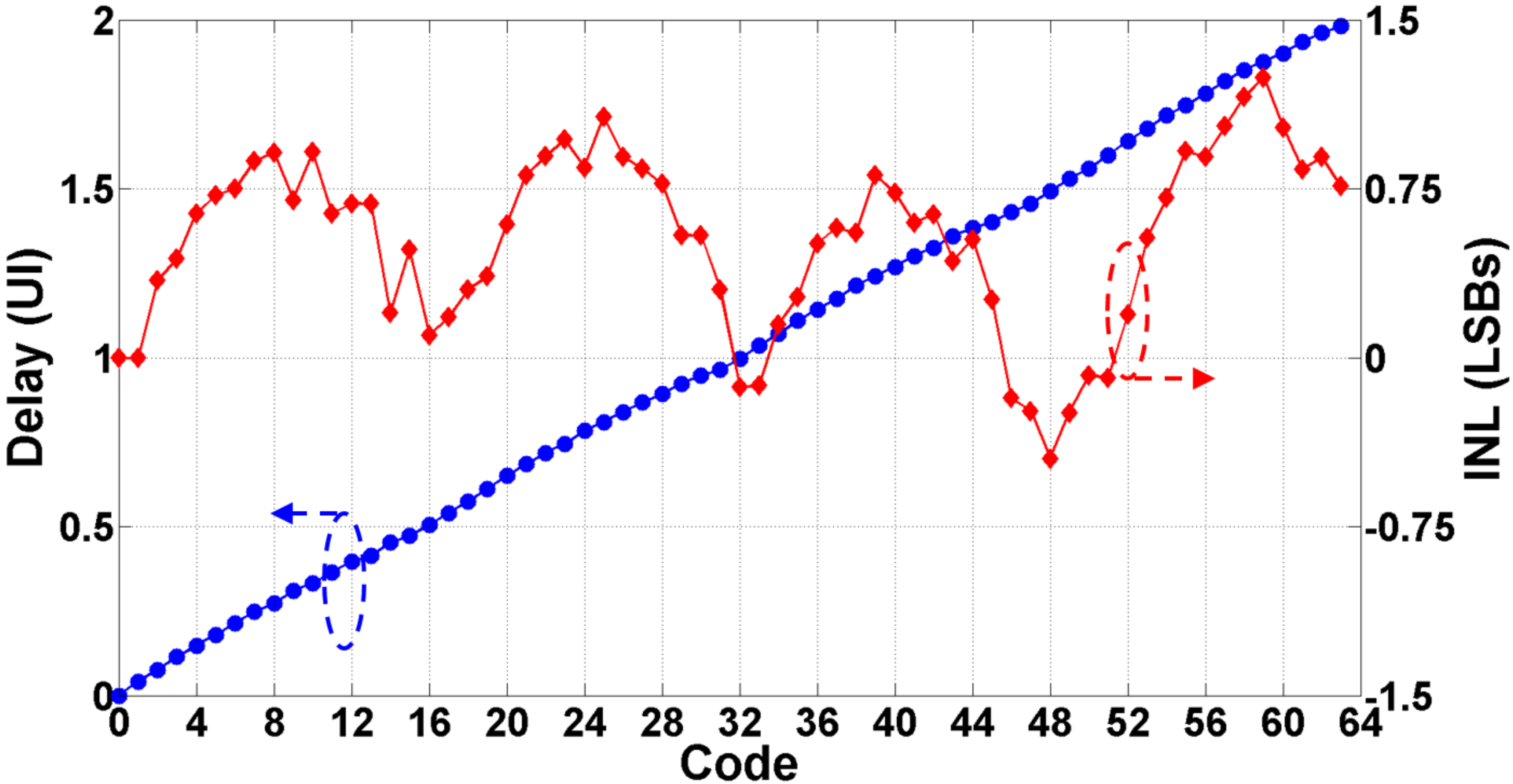


RX Bus Layout



- Fabricated in 32nm SOI CMOS
- Bus-level demonstrator
  - C4-limited test chip area: 2.6 x 3.8 mm²
  - Active circuit area (RX, TX, PLL): 1.46mm²
- Integrated transceiver test chip packaged in an organic SLC
- Modules were mounted to a pin array and plugged into sockets on various test boards

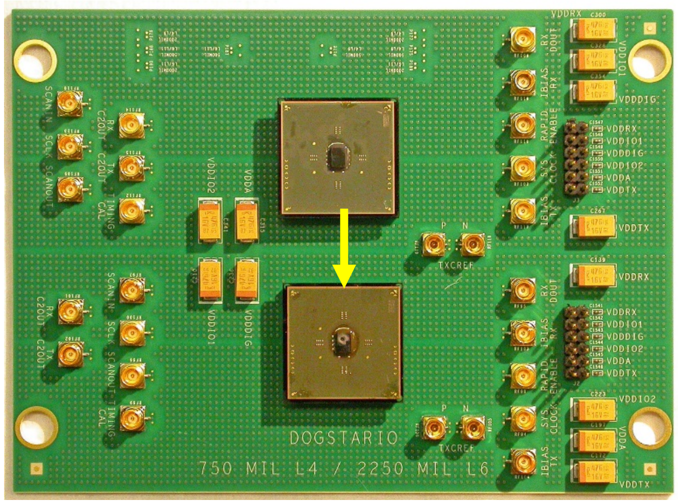
8 GHz Phase Rotator – Measured Performance

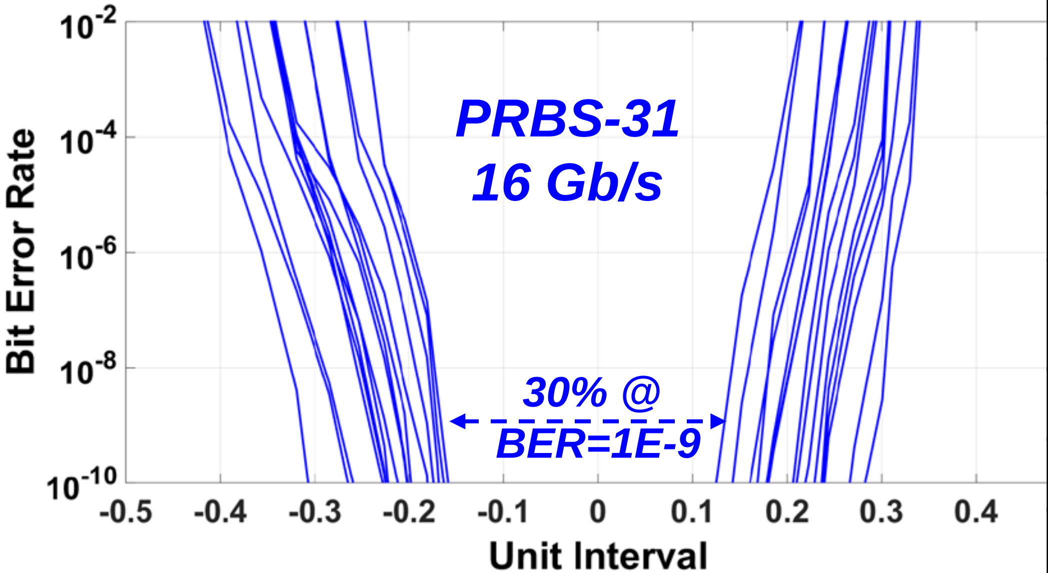


- INL < 1.2 LSBs
  - *Suitable for digital-CDR-based RX in addition to source-synchronous I/O in this work*

Link Measurements: 0.75” USR Link

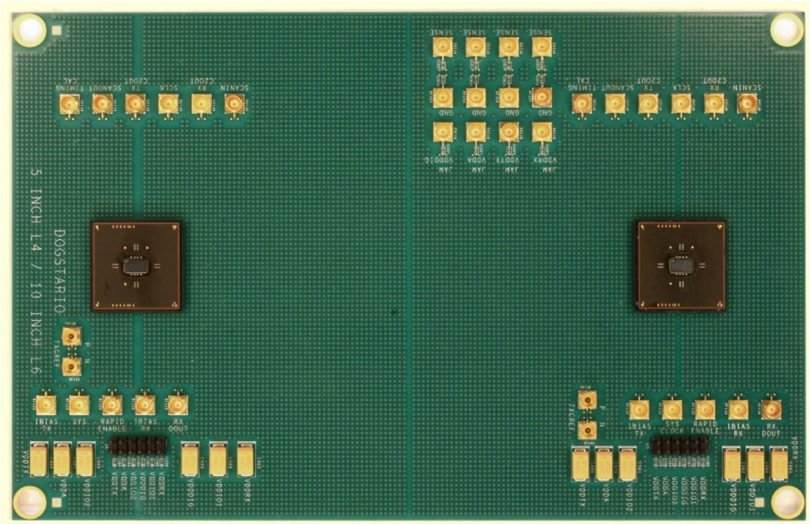
- 16x16 Gb/s links communicating over 0.75” ultra-short reach link on Megtron-6 PCB
  - C4-to-C4 loss of ~6dB at 8 GHz
- Test board & package designed with matched trace lengths to minimize lane-to-lane skew



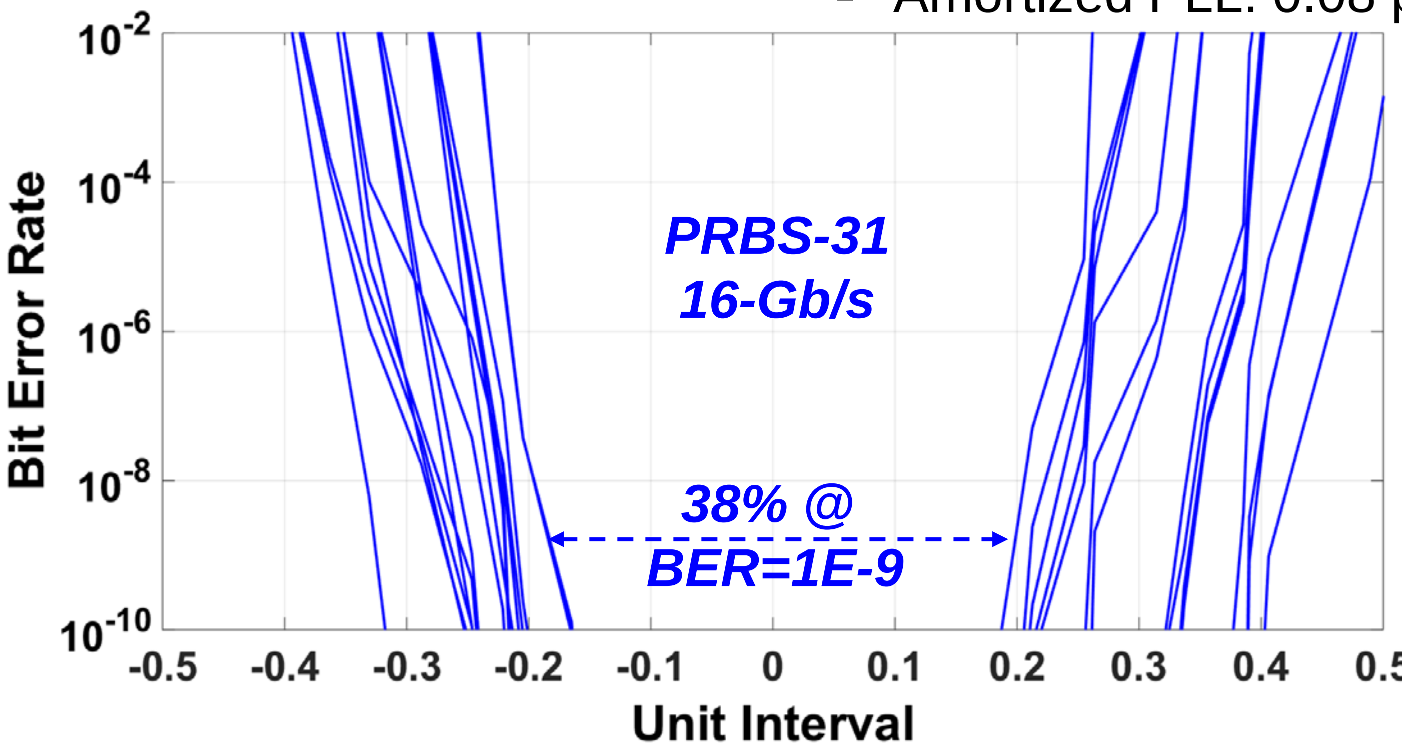


Block	Percentage
TX Data Path/Drivers	24.3%
TX Clock Distribution	3.8%
Forwarded Clock	2.3%
TX total	***30.4%***
RX Data Path	48.6%
RX Phase Recovery and Clock Distribution	18.8 %
RX total	***65.2%***
PLL	***4.4%***
Total	***1.81 pJ/bit***

Link Measurements: 10” Chip-to-Chip Link



- C4-to-C4 loss of 10dB at 8GHz
  - ~200mVppd TX launch amplitude
- Power efficiency: 1.9 pJ/bit
  - TX: 0.67 pJ/bit
  - RX: 1.18 pJ/bit
  - Amortized PLL: 0.08 pJ/bit



Comparison

	Chang et al. VLSI 2008 [7]	Amirkhany et al. ISSCC 2012 [8]	Kaviani et al ISSCC 2012 [9]	Mansuri et al ISSCC 2013 [10]	This Work	
Technology	65nm	40nm	40nm	32nm	32nm SOI	
Per-Lane Data Rate (Gb/s)	16	16	16/20	16	16	
Equalization	TX FFE-5, RX CTLE (CTRL)	Coded differential signaling	TX FFE-2 RX CTLE RX DFE-1	TX FFE-3 RX CTLE	RX CTLE	
Channel Loss (dB)	15	10	15	11	6.2	10
Power Efficiency (pJ/bit)	13 (CTRL)/ 8 (DRAM)	4.1	5.3/6.1	2.6	1.8	1.9

Conclusions

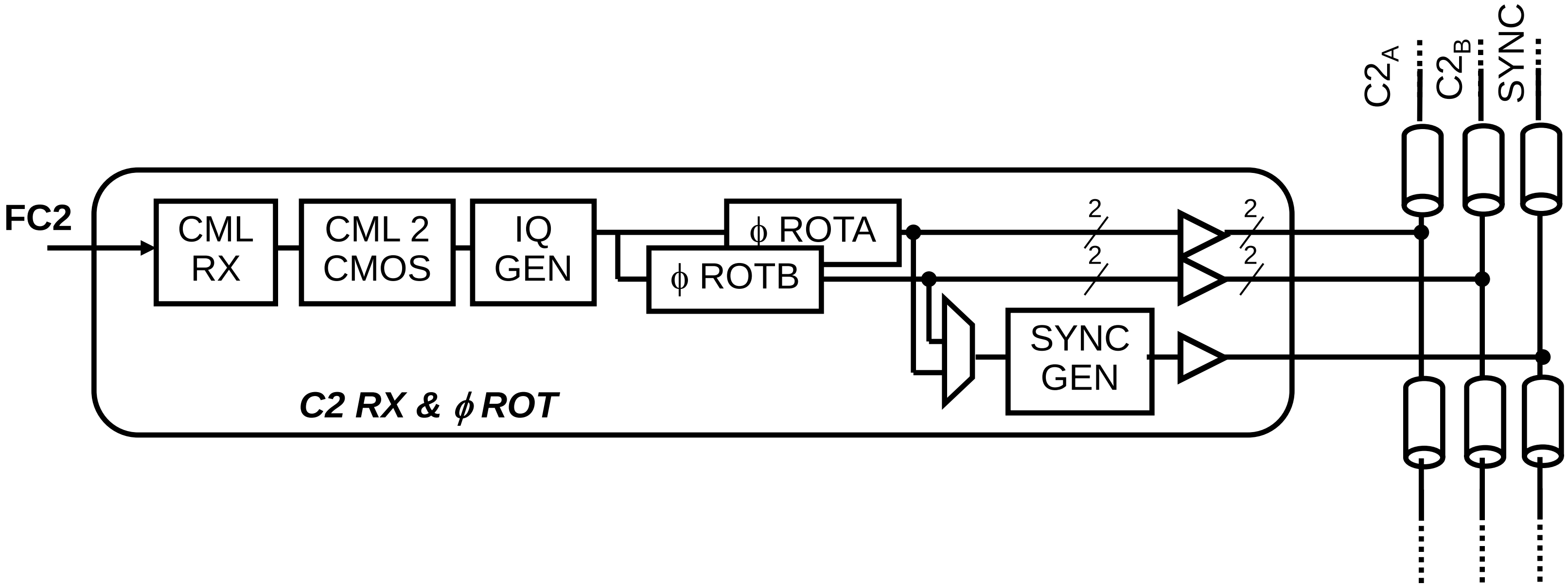
- **Source-synchronous I/O architecture described with new receiver redundancy technique**
  - ***Enables periodic link recalibration with reduced power and area overhead***

- **Described 32nm SOI CMOS implementation of an I/O test chip employing this architecture**
  - ***Achieves better than 2pJ/bit power efficiency operating over interconnects up to 10”***

- **Demonstrates suitability of the I/O architecture for high-density, power-critical USR chip-to-chip applications**

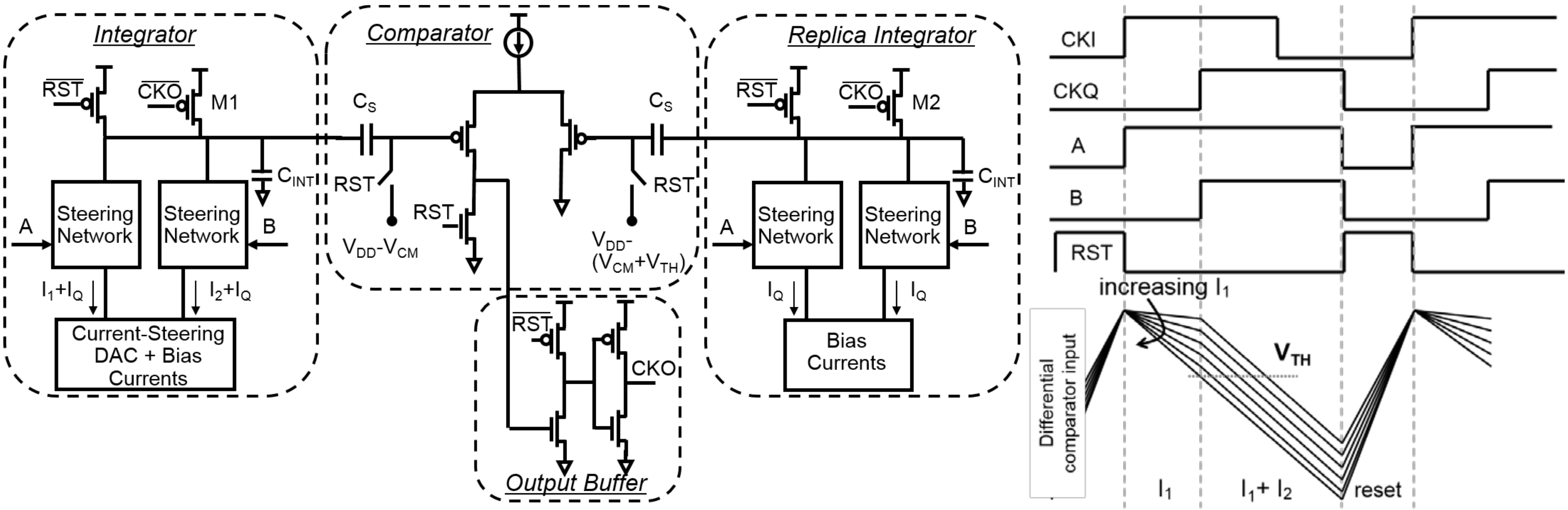
Backup

Forwarded Clock Receiver



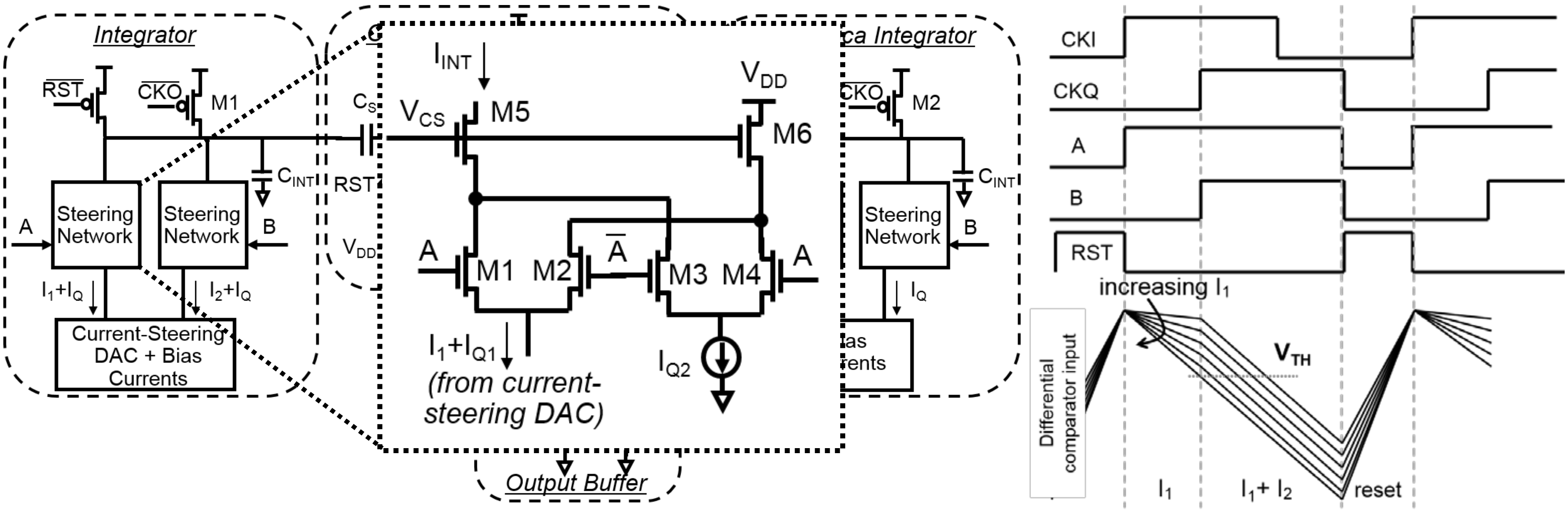
- CML clock receiver to detect low-swing half-rate forwarded clock from TX
- CML to CMOS conversion with duty cycle correction
- Phase rotator based on current-integrating phase interpolators for global phase adjustment
- Synchronization pulse distributed to all RX

Current-Integrating Phase Interpolator (Implementation)



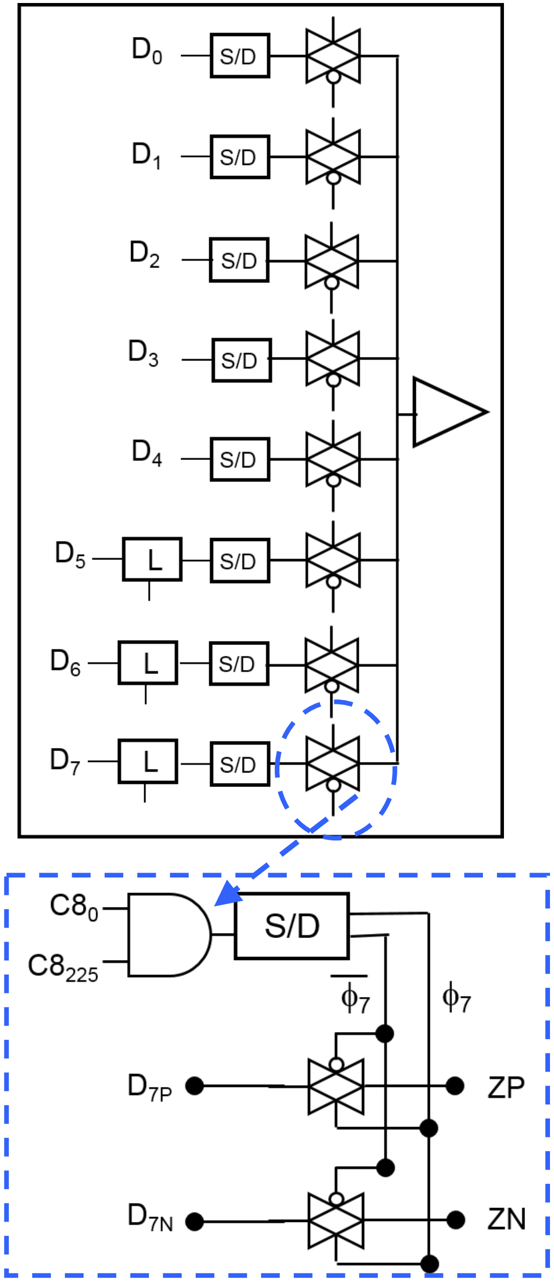
- Signals A & B from pulse generator trigger current integration events within integrator 'Steering Networks'

Current-Integrating Phase Interpolator (Implementation)

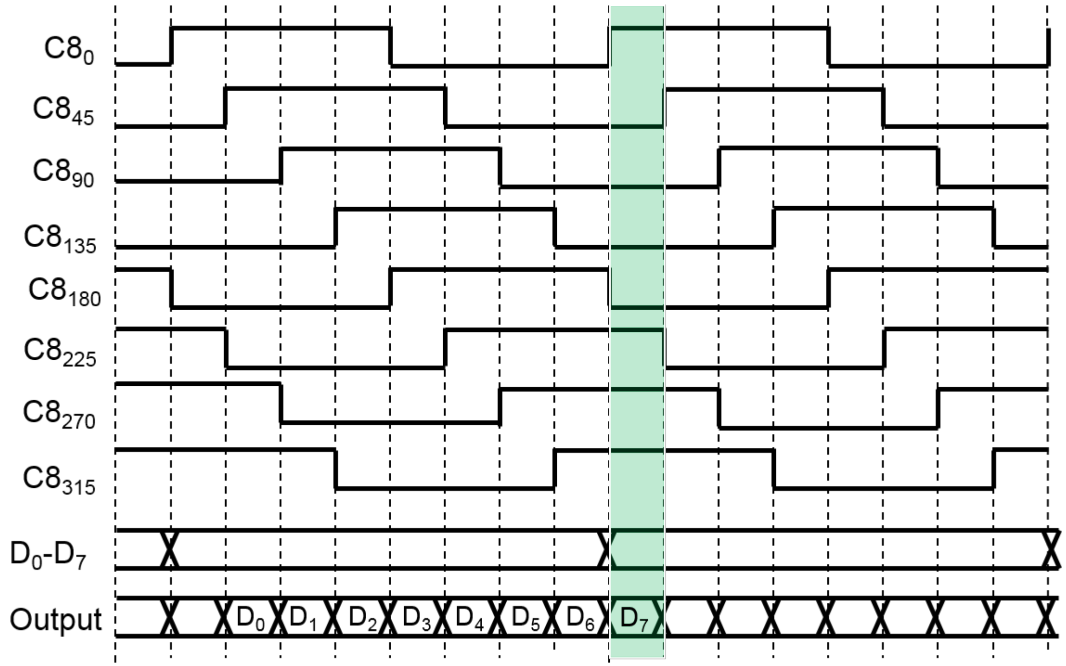


- Cascode device M5 isolates integration node from parasitics in the rest of the switching network
- Inclusion of small quiescent current I_{Q1} avoids nonlinearities with the integration of near-zero currents for certain DAC codes
- I_{Q2} provides small bias to the cascode during reset, to improve transient performance

Transmit 8:1 Serializer



Timing Diagram



- Transmission-gate serializer clocked with 8 eighth-rate (C8) clock phases
  - One of eight paths enabled for 1 UI at a time
  - Lower latency as compared to tree-like serializer structures