



ISSCC 2023

SESSION 28

High-Density Memories and High-Speed Interfaces

A 1.67-Tb, 5b/Cell Flash Memory Fabricated in 192-Layer Floating Gate 3D-NAND Technology and Featuring a 23.3Gb/mm² Bit Density

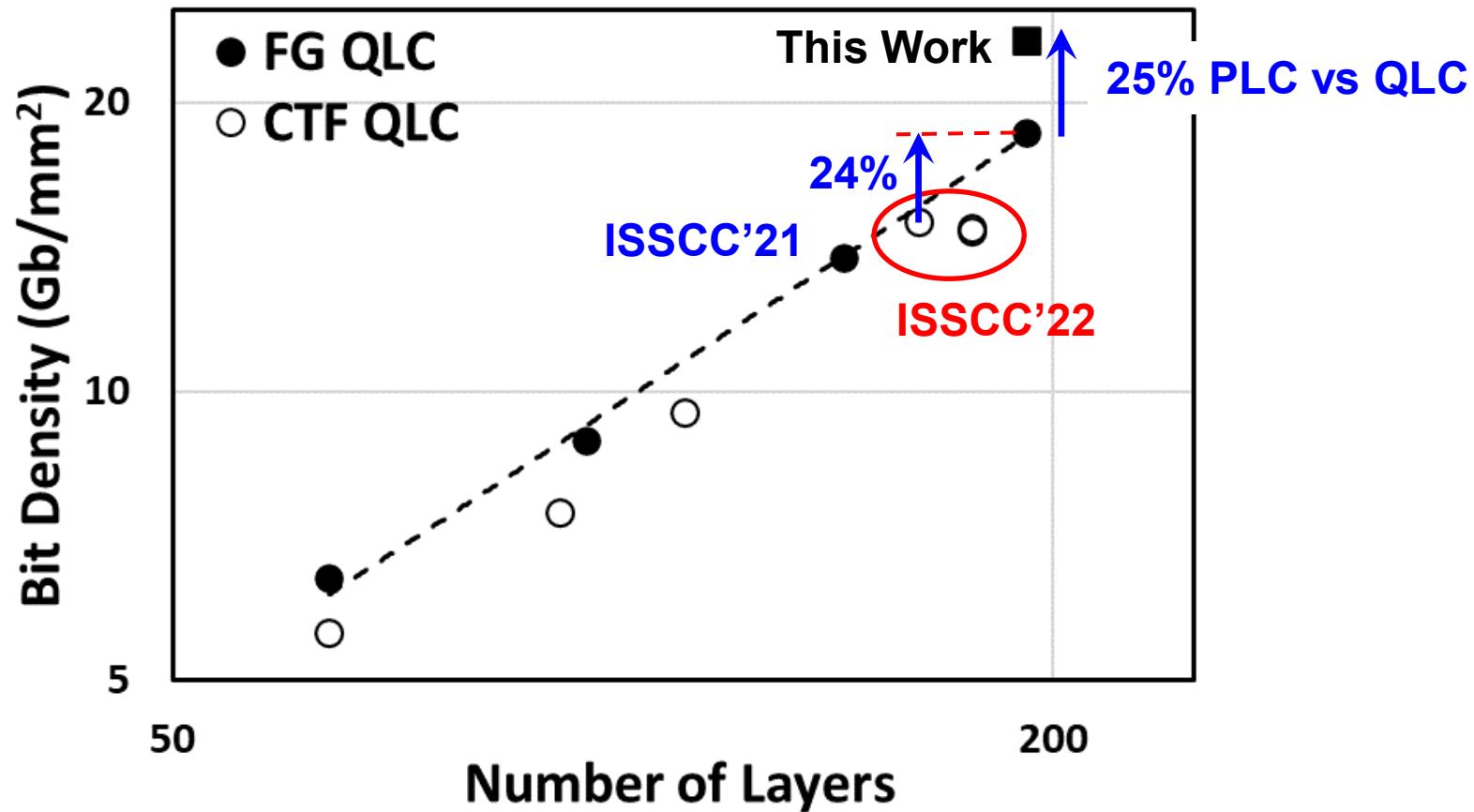
A. Khakifirooz, E. Anaya, S. Balasubrahmanyam, G. Bennett, D. Castro, J. Egler, K. Fan, R. Ferdous, K. Ganapathi, O. Guzman, C.W. Ha, R. Haque, V. Harish, M. Jalalifar, O.W. Jungroth, S.-T. Kang, G. Karbasian, J.-Y. Kim, S. Li, A.S. Madraswala, S. Maddukuri, A. Mohammed, S. Mookiah, S. Nagabhushan, B. Ngo, D. Patel, S.K. Poosarla, N.V. Prabhu, C. Quiroga, S. Rajwade, A. Rahman, J. Shah, R.S. Shenoy, E. Tachie Menson, A. Tankasala, S.K. Thirumala, S. Upadhyay, K. Upadhyayula, A. Velasco, N.K.B. Vemula, B. Venkataramaiah, J. Zhou, B.M. Pathak, P. Kalavade

Intel Corporation, CA, USA

Outline

- **Motivation**
- **Challenges of Increasing Bit Density**
- **Fast Soft-Bit Read (FSBR) Algorithm**
- **Read Calibration Algorithm**
- **Reverse Read Waveform**
- **Program Suspend and Resume**
- **Summary and Conclusions**

Motivation



Build upon successful commercialization of QLC and accelerate bit density increase to extend NAND storage to cost-sensitive markets.

Choice of Program Algorithm

QLC CTF

Coarse-Fine Algo

Balanced Gray code

SLC cache costs 12%

QLC FG [1]

4-16 Algo

(Max t_R / Avg t_R) ~ 1.5

No need to SLC cache

PLC (This Work)

Coarse-Fine Algo

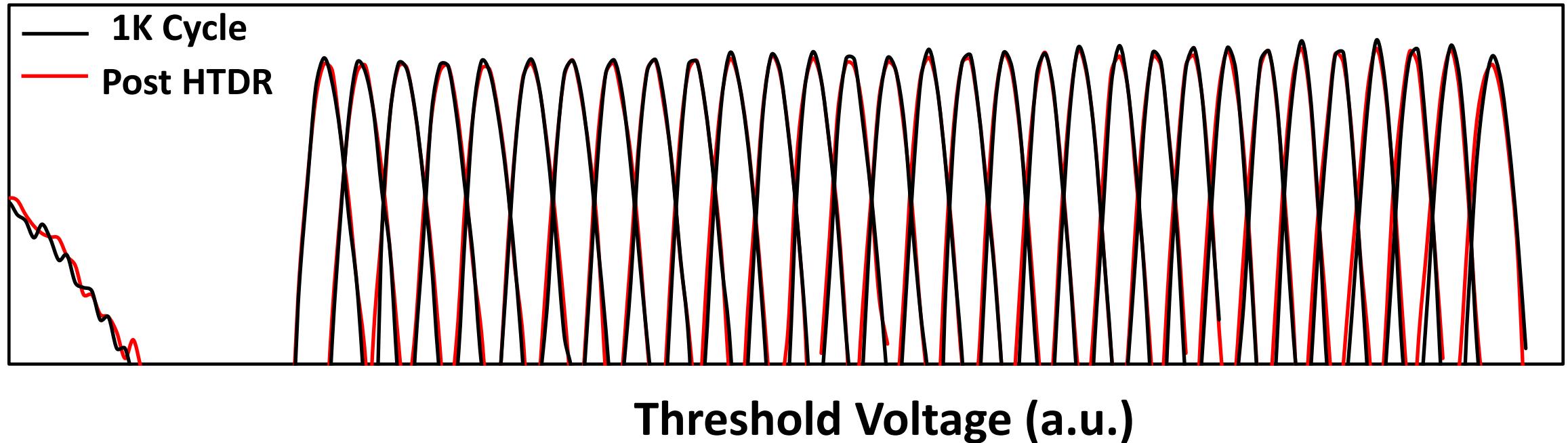
Balanced Gray code

SLC cost is kept at 2%

- We used coarse-fine algorithm to allow balanced Gray code.
- SLC endurance was increased to 250K to limit the cost of SLC cache to 2%.

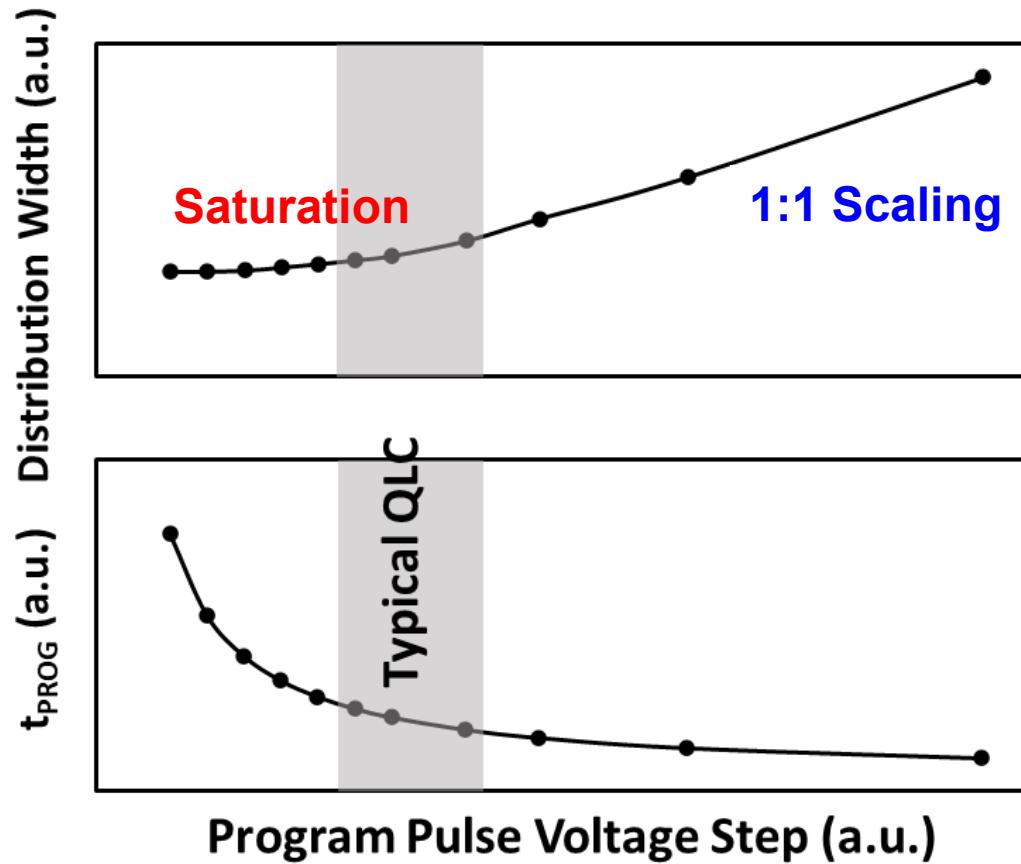
$$\text{SLC Cost} = 5 \frac{\text{PLC Endurance}}{\text{SLC Endurance}} = 5 \frac{1 \text{ K}}{250 \text{ K}}$$

Enabling PLC with Floating Gate Technology



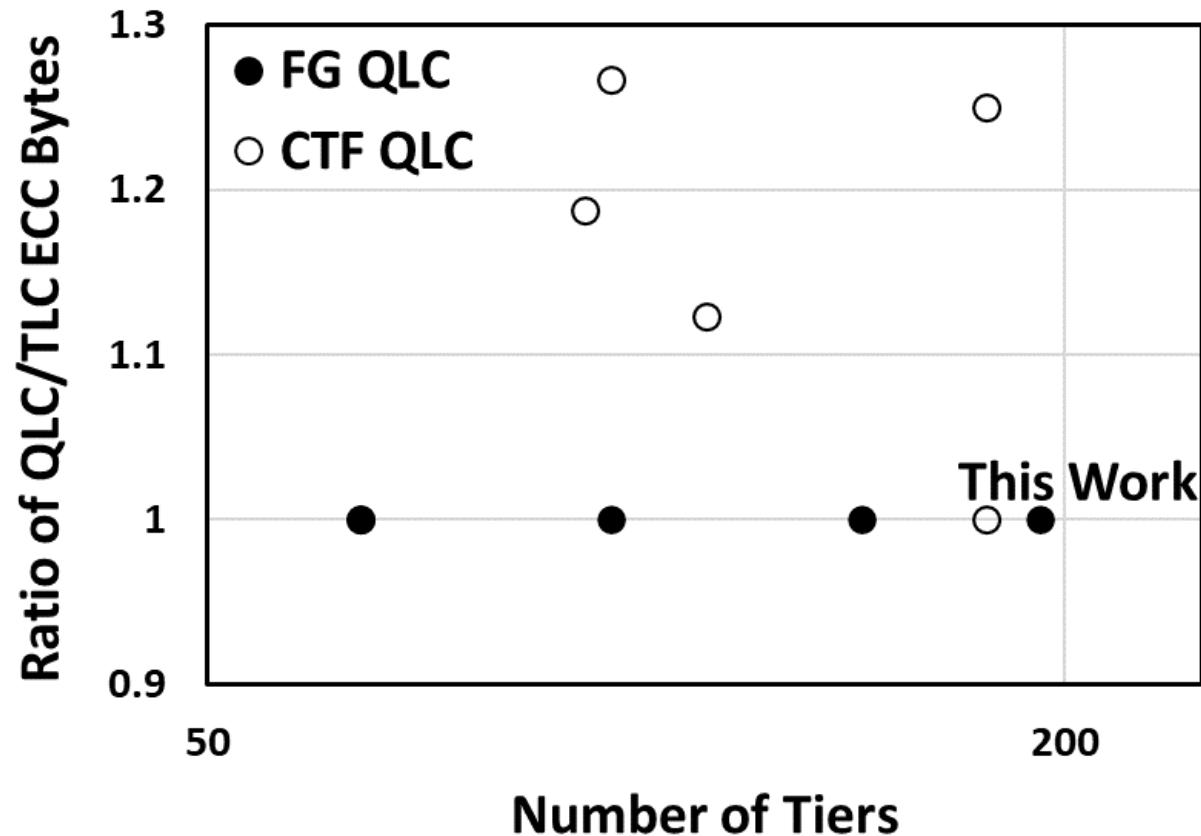
Intrinsic resilience of the FG technology to charge loss (isolated charge storage nodes) facilitates PLC implementation.

Challenges of PLC V_T Placement



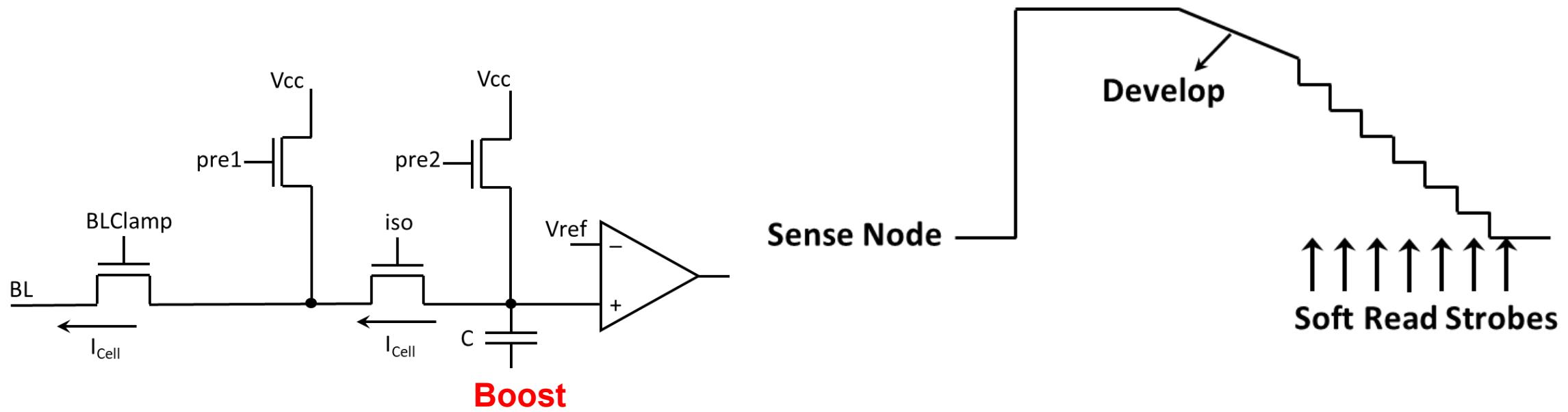
- Reducing program gate step beyond QLC is inefficient due to RTN.
- Need to increase error correction capability.

Increasing Error Correction Code Capability



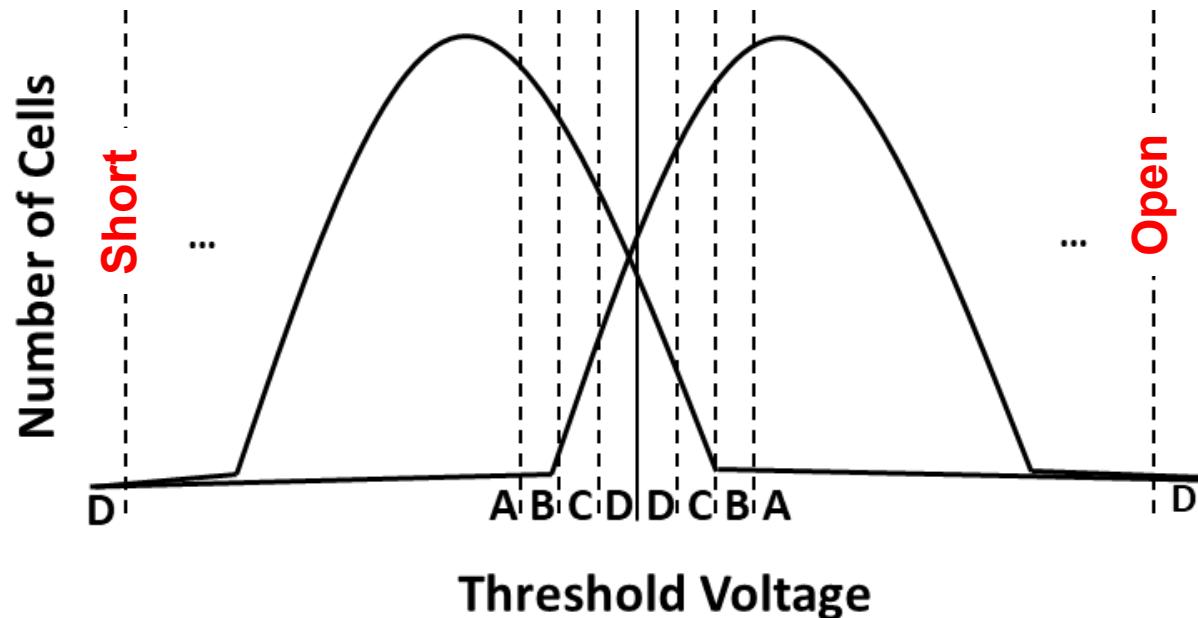
- Most QLC technologies used higher number of ECC bytes vs TLC.
- To avoid cost increase, we maintained the number of ECC bytes.

Fast Soft-Bit Read Algorithm



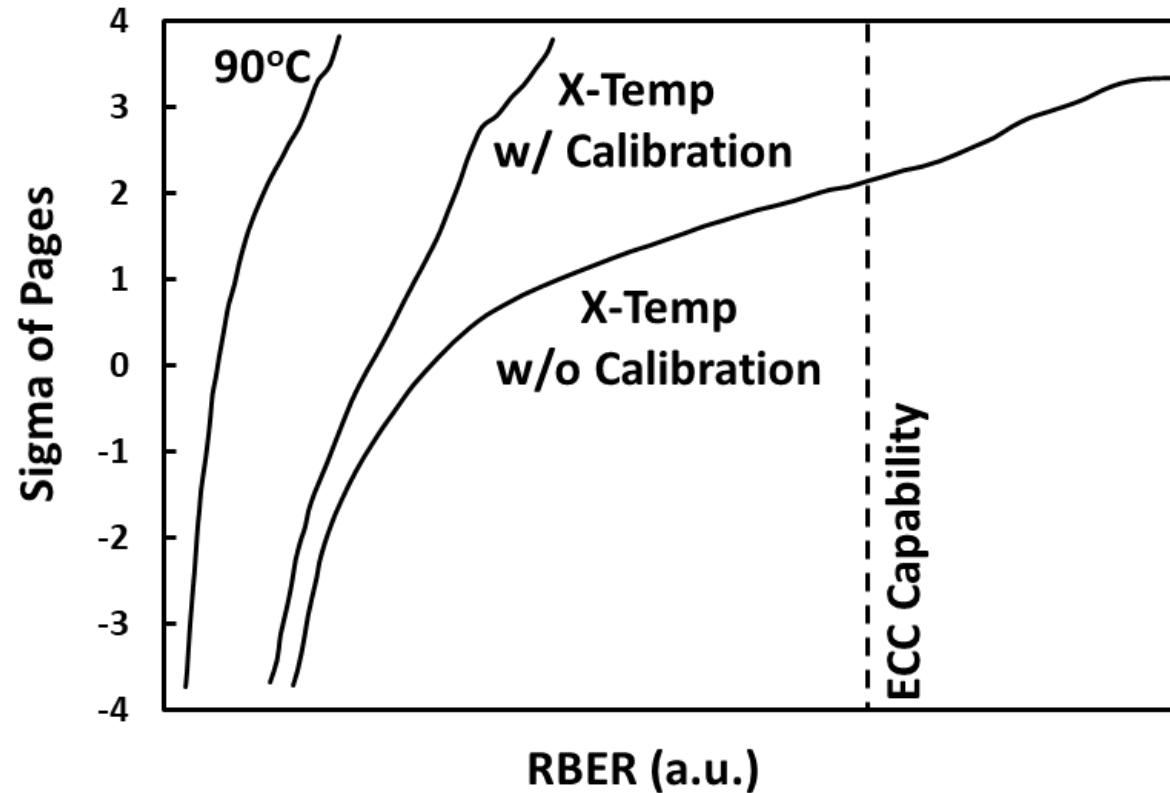
- **Fast SBR uses boost voltage modulation **after** sense capacitor is isolated from BLs, to sense the cells at different current.**
- **Cells are grouped into 4 buckets from highest to lowest confidence.**

FSBR with Unrepaired Defective BLs



- To further reduce cost, we reduced number of redundant columns by >70%.
- Unrepaired BLs are allowed as long as they are way below ECC limit.
- Special open/short detection was added to the read algorithm to place these BLs in the lowest confidence bucket.

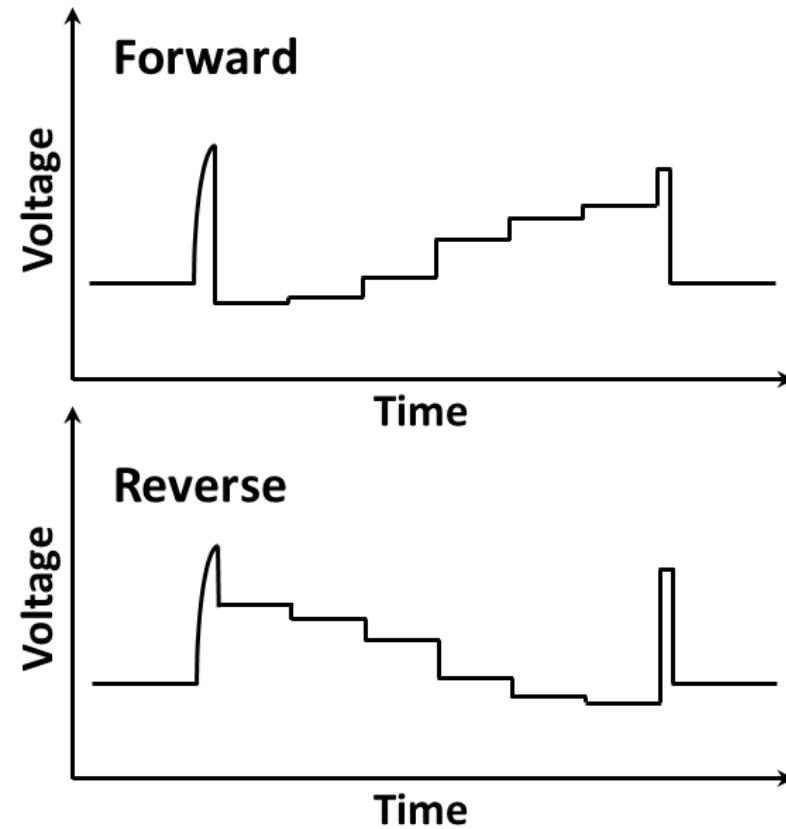
Fast Auto-Read Calibration



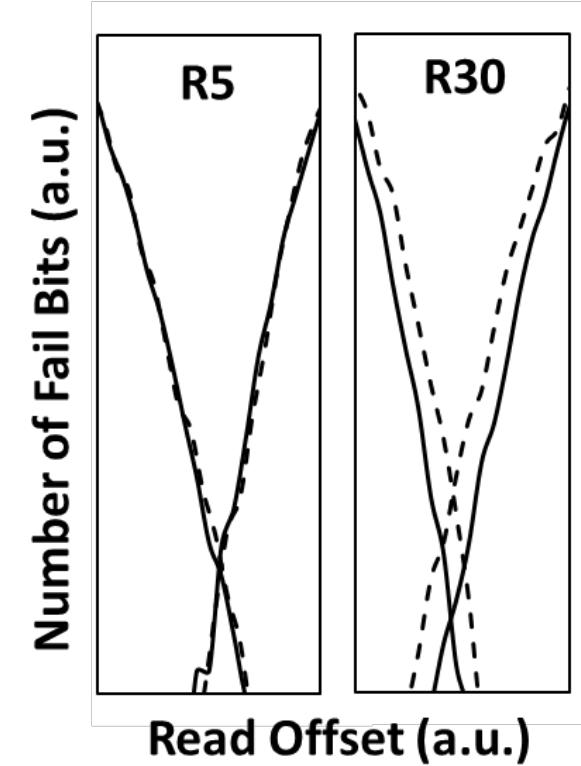
- **Fast ARC algorithm uses 5-strobe boost modulation to calculate optimum read level.**
- **This is more accurate than 3-strobe algorithm presented in [1].**

Reverse Read for Improved Read Margin

Cells in depletion prior to sense

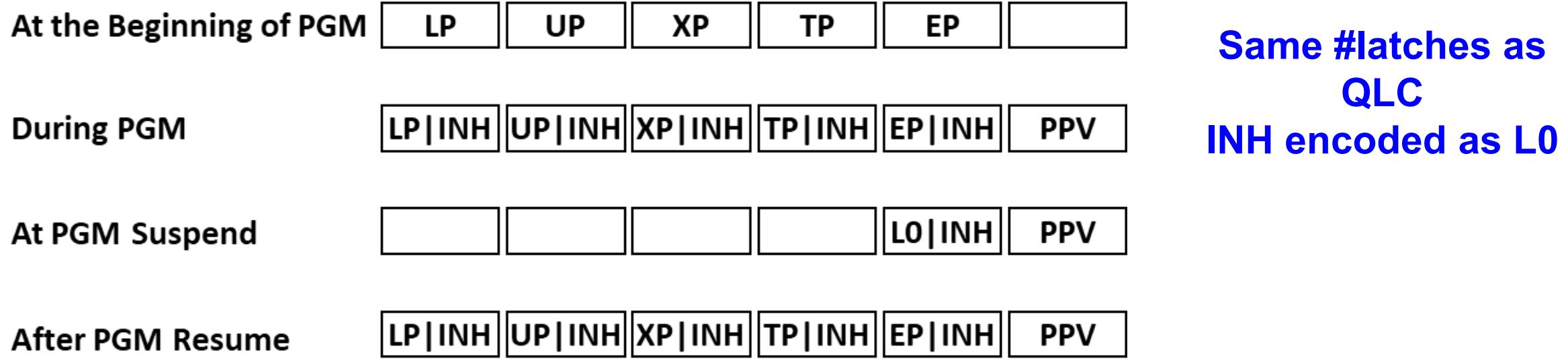


Cells in inversion prior to sense



- Reverse read waveform is used to increase read margin for higher levels by keeping cells in inversion prior to sensing.

Program Suspend and Resume Algorithm

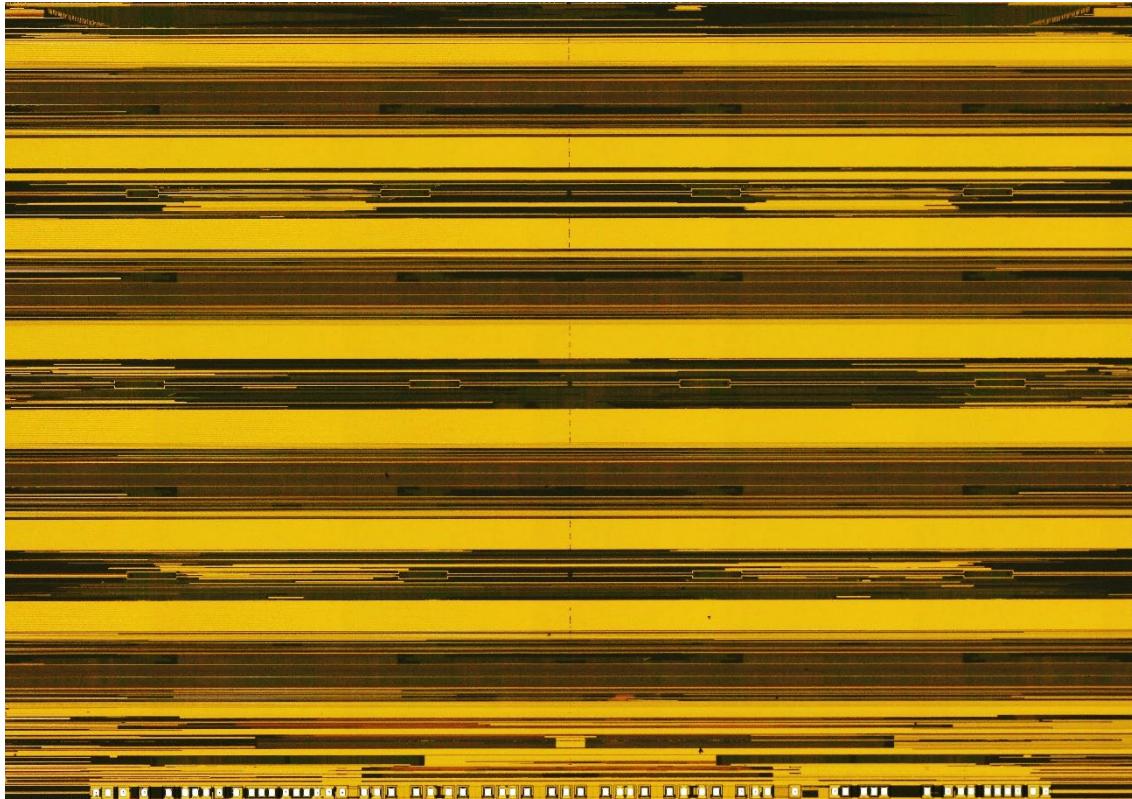


- Need at least 3 latches to support FSBR and FARC during program suspend.
- Leveraging user data in SLC cache to free-up the latches at suspend.
- Data is reconstructed at program resume.

Conclusions

- We reported the **first 5 bit/cell NAND flash, delivering a record bit density of 23.3Gb/mm².**
- The chip can operate in QLC or TLC modes delivering 24% higher bit density compared to the best reported QLC densities.
- A fast soft-bit read algorithm was implemented to augment ECC capability and is capable of handling defective BLs.
- A fast read calibration algorithm was implemented to calculate optimum read level.
- Reverse read waveform was used to increase read margin.
- Program suspend and resume algorithms compatible with FSBR and FARC were implemented.

Key Features



Number of Layers	192
Capacity	1.67 Tb
Number of Planes	4
Program Time (μs)	5500
Read Time (μs)	354
Endurance (P/E Cycle)	1K
Die Size (mm²)	73.3
Bit Density (Gb/mm²)	23.3
I/O Rate (MT/s)	1600

28.1: A 1.67-Tb, 5b/Cell Flash Memory Fabricated in 192-Layer Floating Gate 3D-NAND Technology and Featuring a 23.3Gb/mm² Bit Density

A High-Performance 1-Tb 3b/Cell 3D-NAND Flash with a 194MB/s Write Throughput on over 300 Layers

Byungryul Kim, Seungpil Lee, Beomseok Hah, Kangwoo Park, Yongsoon Park, Kangwook Jo, Yujong Noh, Hyeoncheon Seol, Hyunsoo Lee, Jaehyeon Shin, Seongjin Choi, Youngdon Jung, Sungho Ahn, Yonghun Park, Sujeong Oh, Myungsu Kim, Seonguk Kim, Hyunwook Park, Taeho Lee, Haeun Won, Minsung Kim, Cheulhee Koo, Yeonjoo Choi, Suyoung Choi, Sechun Park, Dongkyu Youn, Junyoun Lim, Wonsun Park, hwang Hur, Kichang Kwean, Hongsok Choi, Woopyo Jeong, Sungyong Chung, Jungdal Choi, Seonyong Cha



SK hynix, Icheon, Korea

Self Introduction



Byungryul Kim

- Biography:

- **B.S degree from Sogang University in 2004**
- **Have been with SK Hynix since 2004**
- **Past : 2D MLC design, 3D MLC/TLC design & verification**
- **Present: a design team leader of TLC project in SK hynix**
- **My interests are to develop new features to enhance performance of NAND Flash memory and to apply them efficiently by consuming minimum area**

Outline

■ New Technologies

- Triple-verify Program (TPGM)
- Adaptive Unselected String Pre-charge (AUSP)
- Programmed Dummy String (PDS)
- All-pass Rising (APR)
- Plane-level Read Retry (PLRR)

■ Key Features

■ Conclusion

Outline

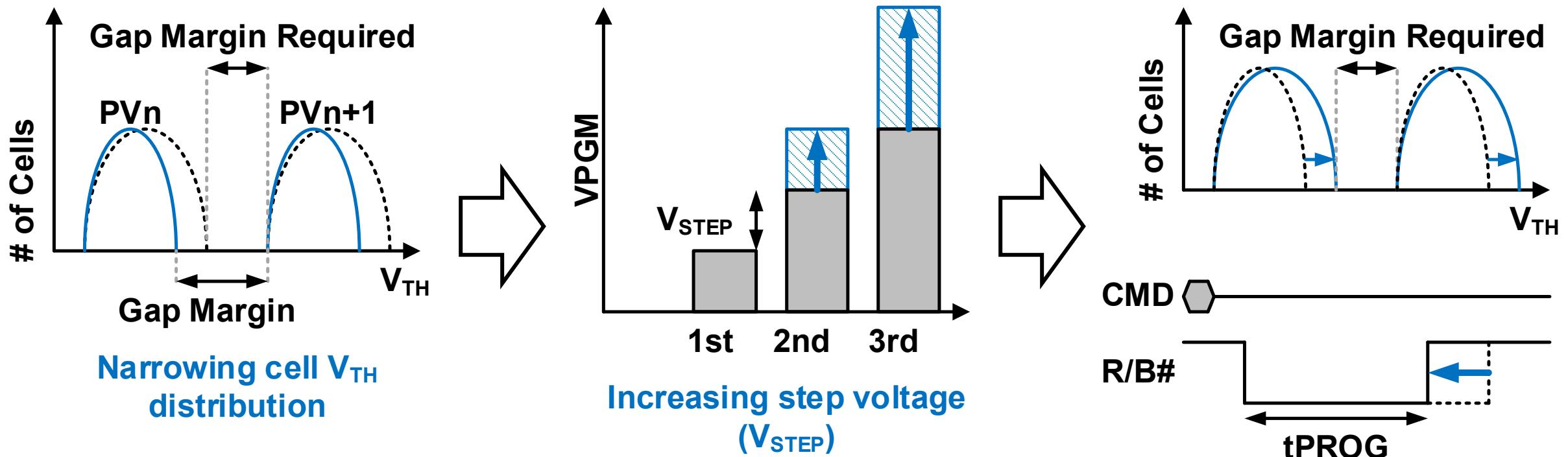
■ New Technologies

- **Triple-verify Program (TPGM)**
- Adaptive Unselected String Pre-charge (AUSP)
- Programmed Dummy String (PDS)
- All-pass Rising (APR)
- Plane-level Read Retry (PLRR)

■ Key Features

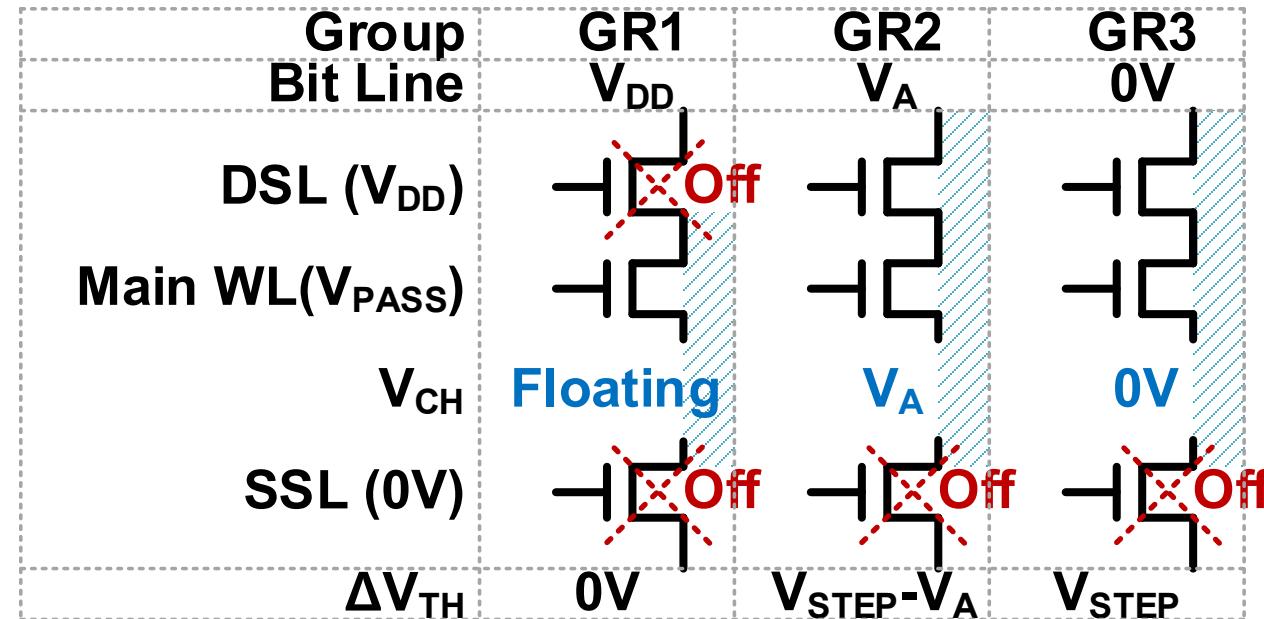
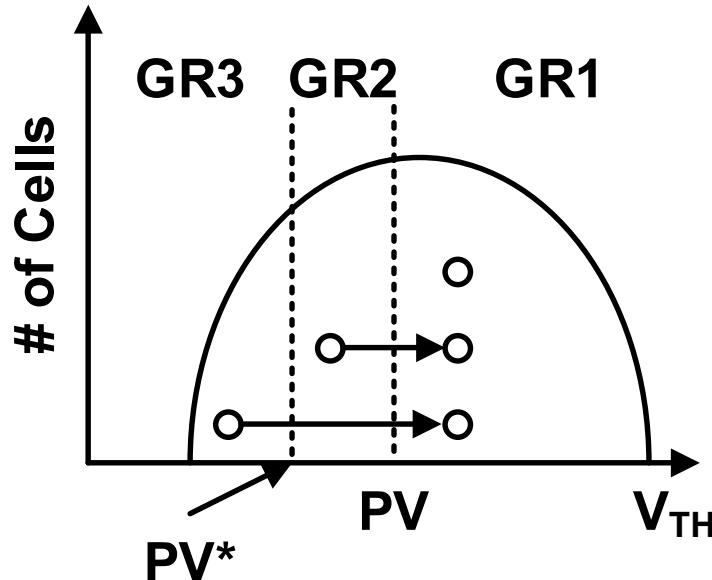
■ Conclusion

PV Distribution vs. tPROG



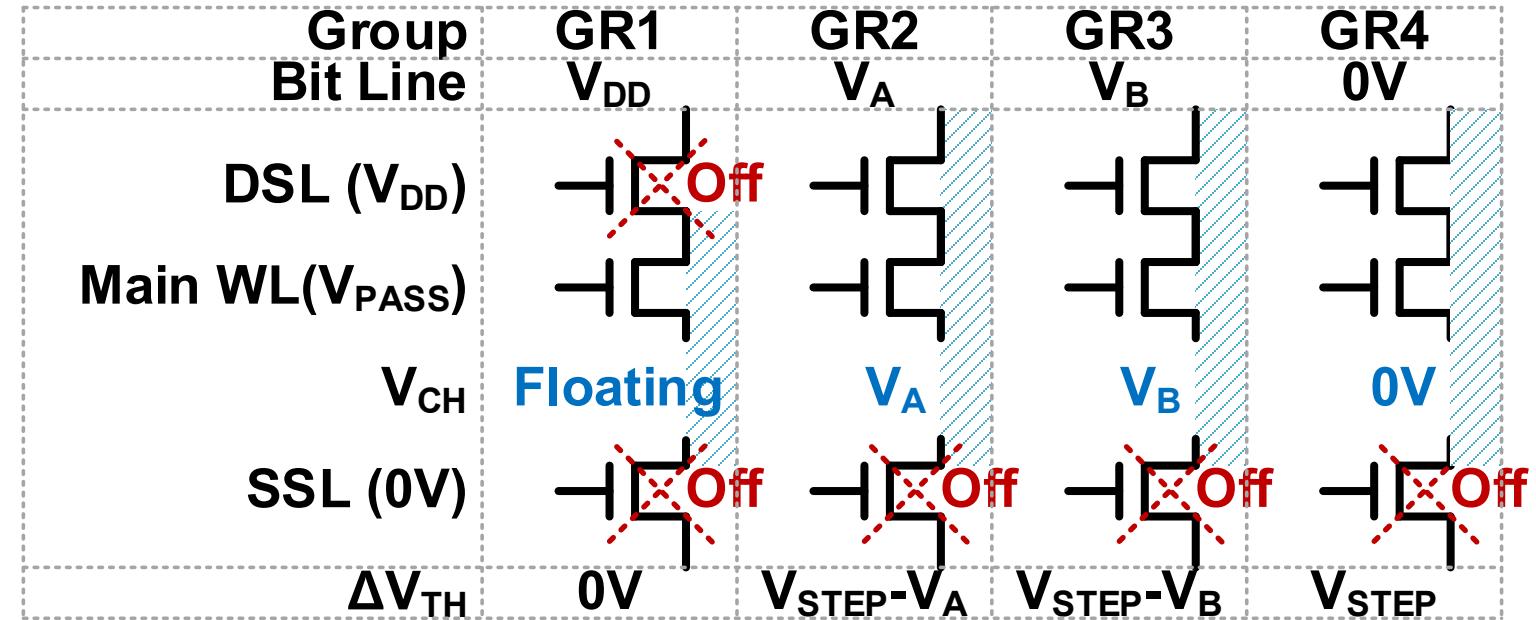
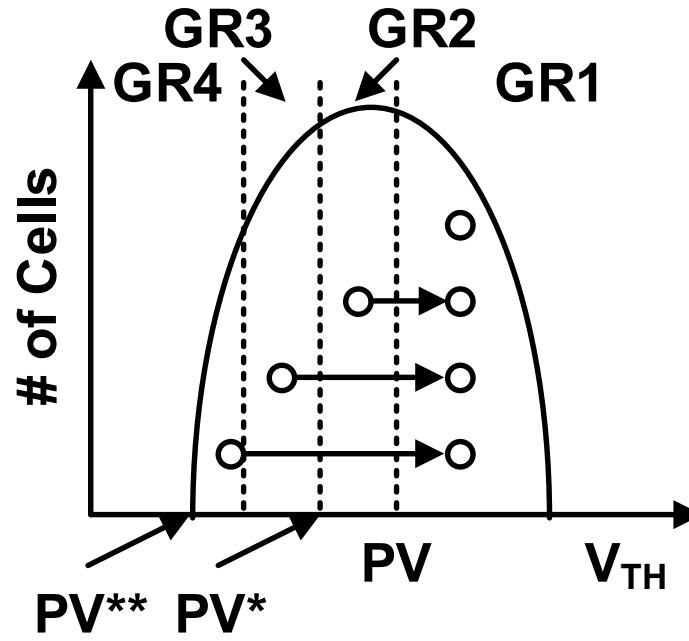
- Narrowing cell V_{TH} distribution enhances Gap Margin
- Larger V_{STEP} makes cell V_{TH} distribution wider and tPROG shorter

Double-verify Program (DPGM)



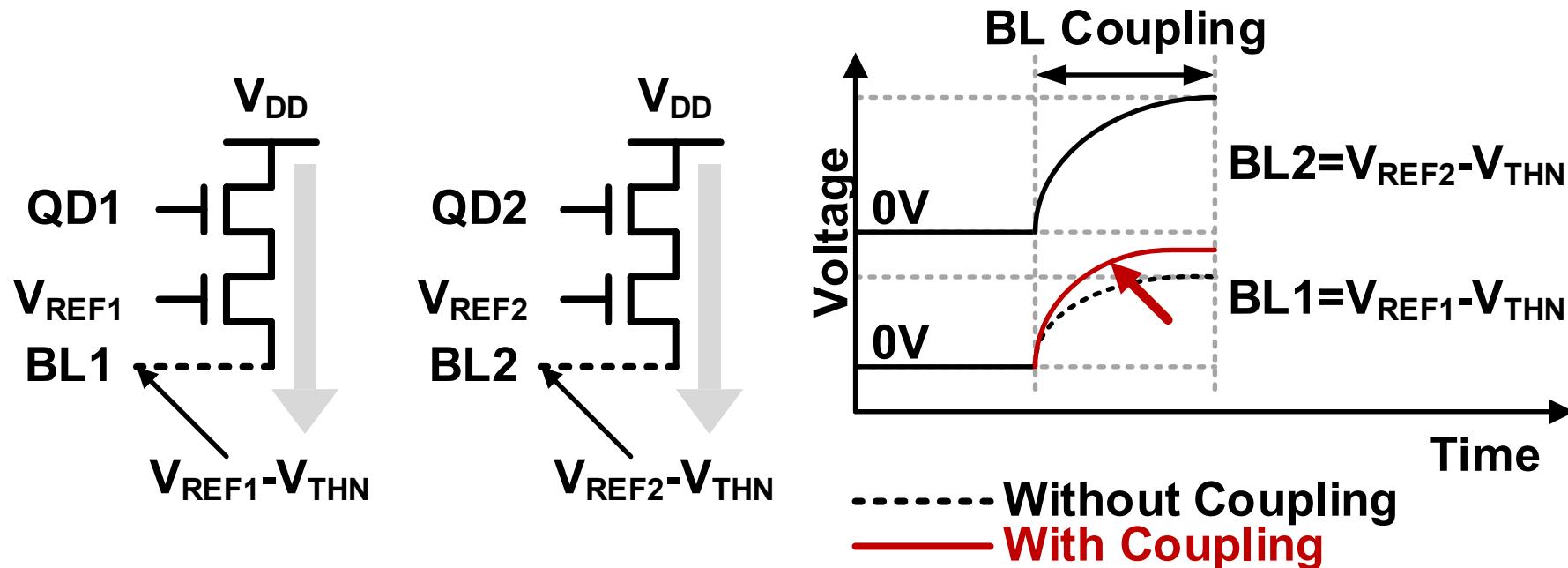
- GR1 BLs= V_{DD} → Channel is Isolated & Not Programmed
- GR2 BLs= V_A → $\Delta V_{TH} = V_{STEP} - V_A$
- GR3 BLs=0V → $\Delta V_{TH} = V_{STEP}$

Triple-verify Program (TPGM)



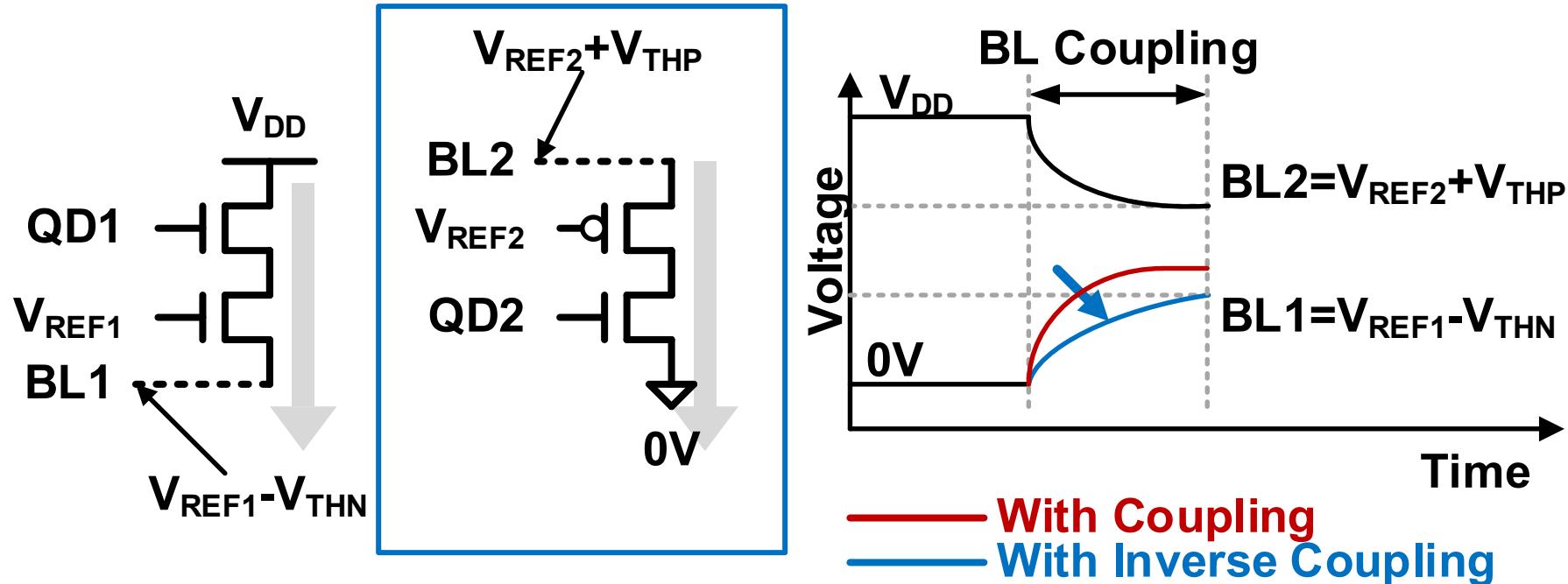
- Adding one more group ($\Delta V_{TH} = V_{STEP} - V_B$, $V_A > V_B$)
- Narrowing the cell threshold voltage (V_{TH}) distribution
- ~10% tPROG reduction with increased V_{STEP}

Conventional BL Pre-charge



- Using two NMOS cascades for $BL_1 = V_B$ and $BL_2 = V_A$ ($V_A > V_B$)
- BL_1 and BL_2 are initially 0V and are set to $V_{REF1} - V_{THN}$ and $V_{REF2} - V_{THN}$
- Coupled by BL_2 , BL_1 exceeds the target level

Counter Driving



- BL1 is initially 0V and is set to $V_{REF1} - V_{THN}$
- BL2 is initially V_{DD} and is discharged to $V_{REF2} + V_{THP}$
- Preventing BL coupling effect by inverse coupling

Outline

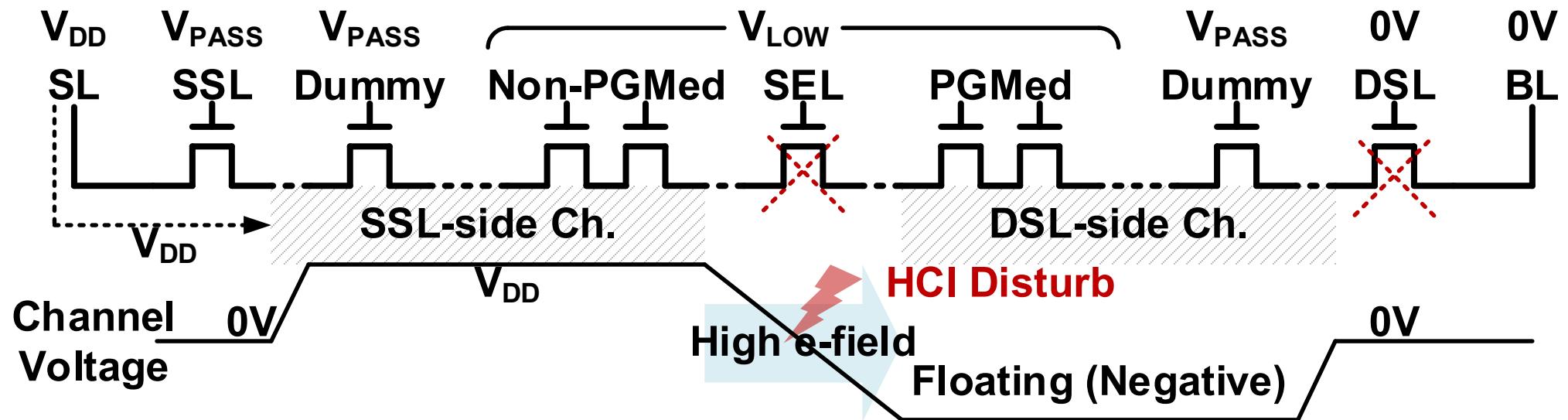
■ New Technologies

- Triple-verify Program (TPGM)
- **Adaptive Unselected String Pre-charge (AUSP)**
- Programmed Dummy String (PDS)
- All-pass Rising (APR)
- Plane-level Read Retry (PLRR)

■ Key Features

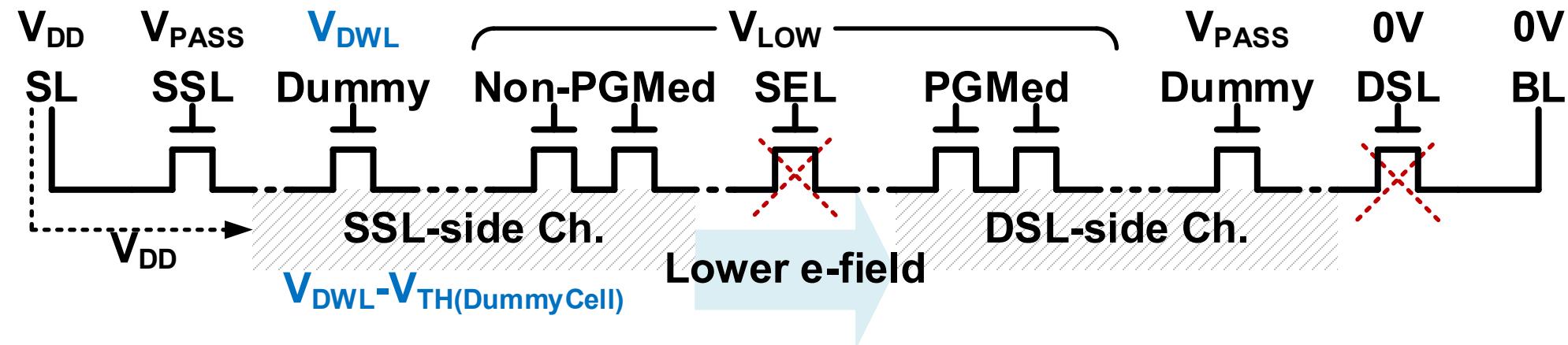
■ Conclusion

Unselected String Pre-charge (USP)



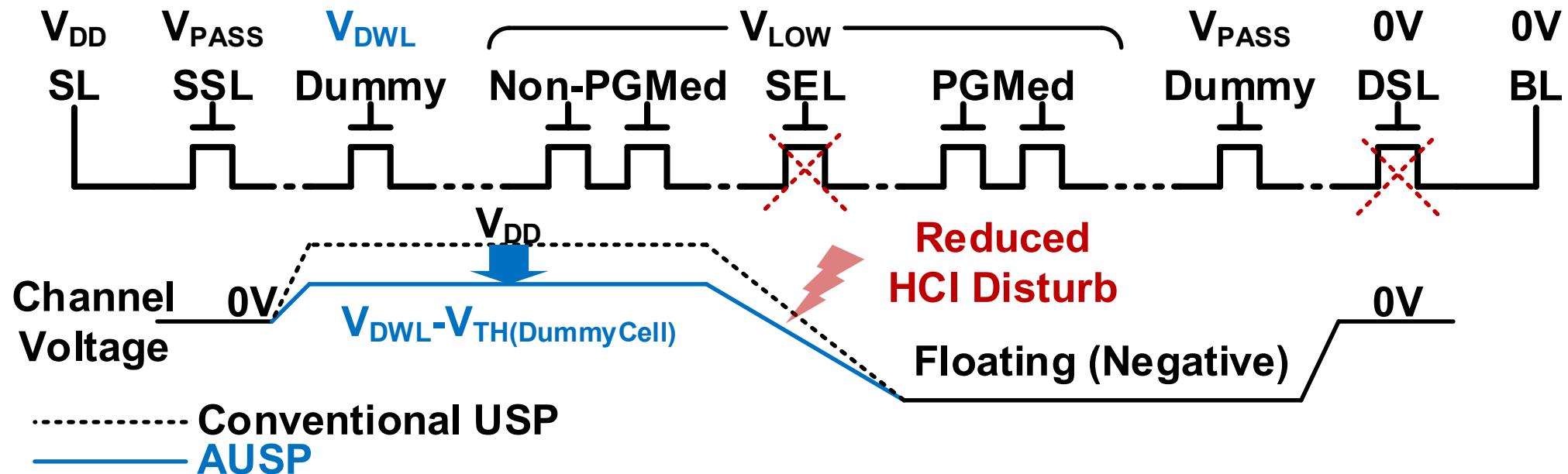
- A program pulse is preceded by an USP (initializing channels)
- USP prevents lack of channel boosting at the program pulse
- SSL-side Channel is pre-charged to V_{DD}
- Hot-carrier injection (HCI) occurs due to high e-field

Adaptive Unselected String Pre-charge (AUSP)



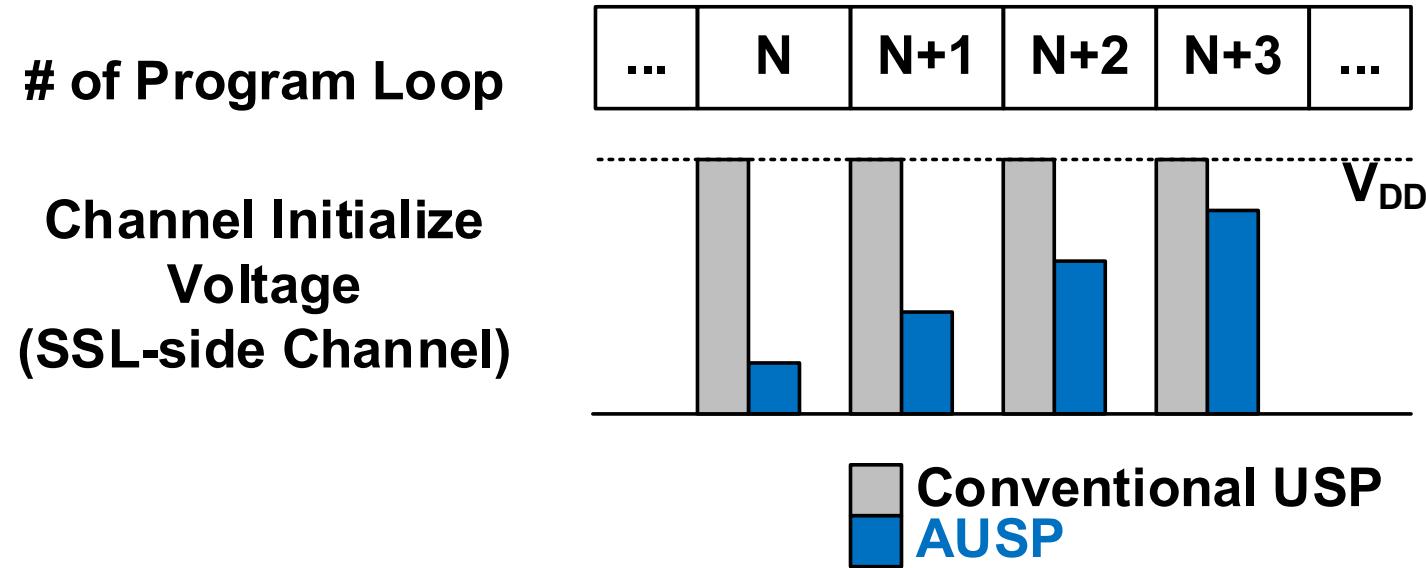
- SSL-side dummy WL is controlled by V_{DWL}
- SSL-side Channel is pre-charged to $V_{DWL} - V_{TH(\text{DummyCell})}$
- HCI disturbances are reduced due to a lower electric field

Channel Initialize Voltage Comparison



- HCI disturbance is produced by channel voltage difference
 - Voltage difference between SSL-side and DSL-side channels
- Channel Voltage is reduced from V_{DD} to $V_{DWL} - V_{TH(DummyCell)}$

Incremental Channel Initialization Voltage



- Channel voltage corresponds to the SSL-side voltage
- Channel voltage can be lowered for lower program loops
 - HCI disturb is reduced further
- Reduced Cell V_{TH} dist. contributes to ~2% tPROG reduction

Outline

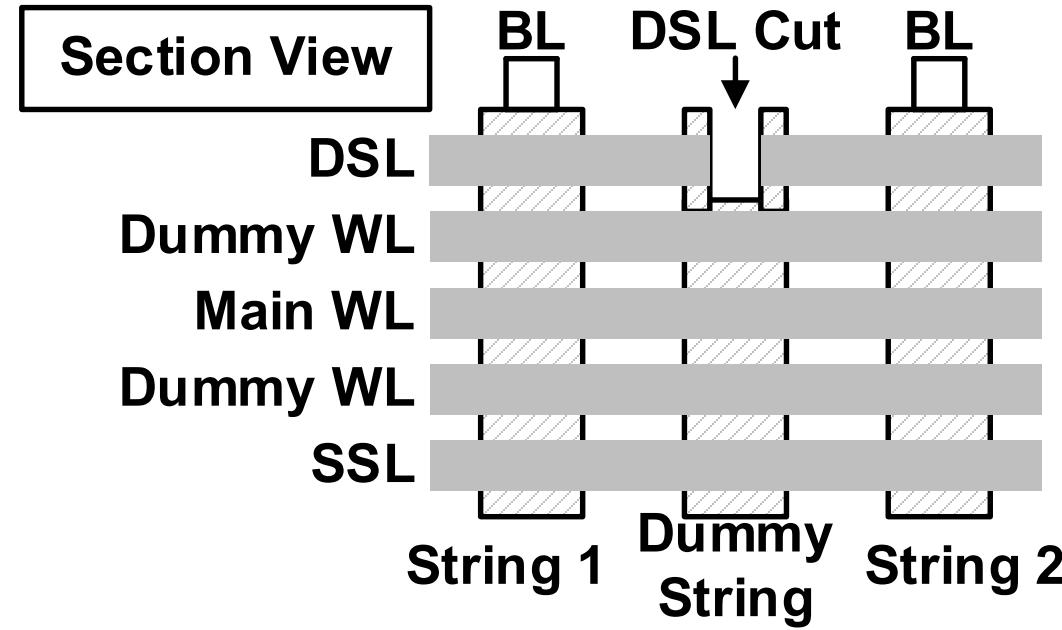
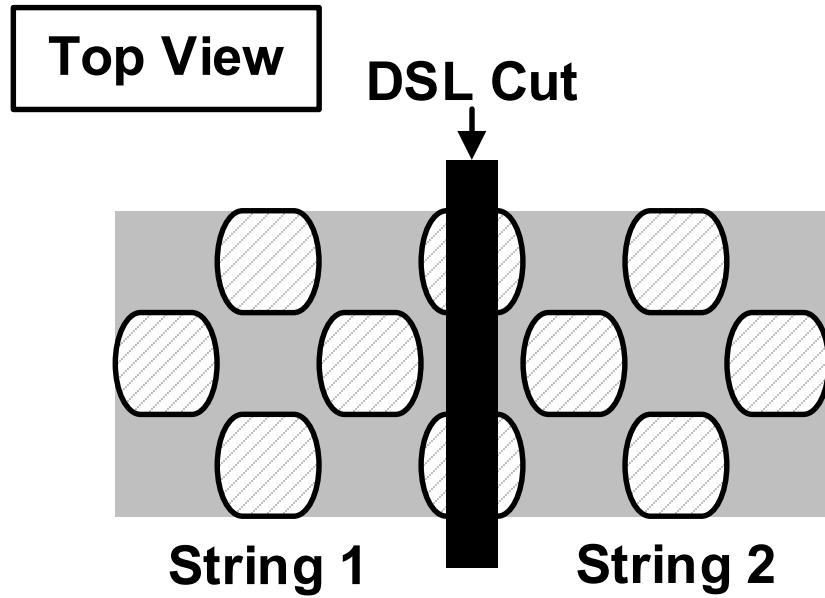
■ New Technologies

- Triple-verify Program (TPGM)
- Adaptive Unselected String Pre-charge (AUSP)
- **Programmed Dummy String (PDS)**
- All-pass Rising (APR)
- Plane-level Read Retry (PLRR)

■ Key Features

■ Conclusion

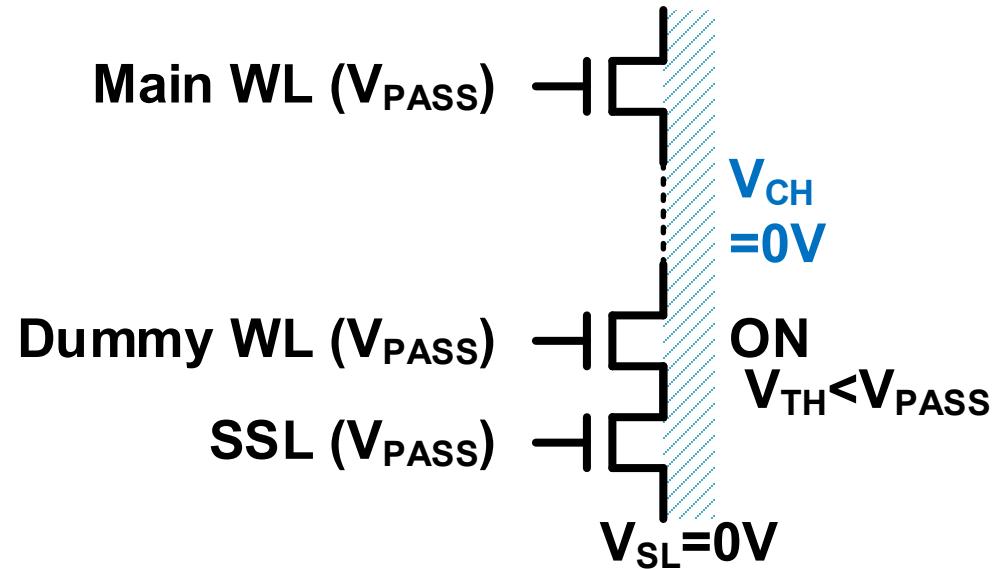
Dummy String



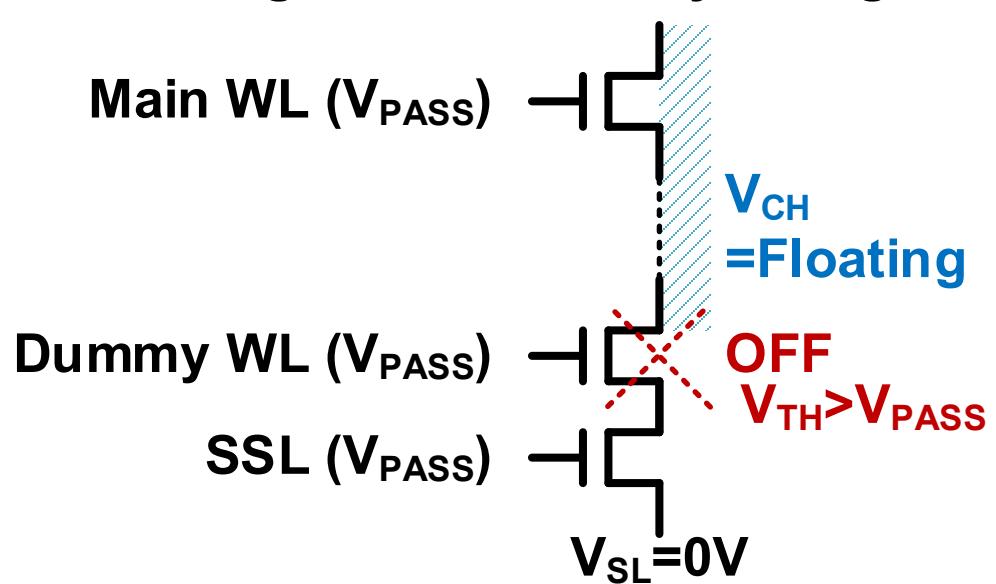
- DSLs are divided by the DSL cut (WLs/SSL are not separated)
- Dummy strings are produced by the DSL cut
- Capacitive load for WL rising/falling is increased

Programmed Dummy String (PDS)

Non-programmed Dummy String



Programmed Dummy String



- Dummy WL Tr. $V_{TH}=1V$, channel voltage becomes 0V
- Dummy WL Tr. is programed ($V_{TH}>V_{PASS}$), channel is floated
- Floating channel reduces cap. load and affects to WL rising

Outline

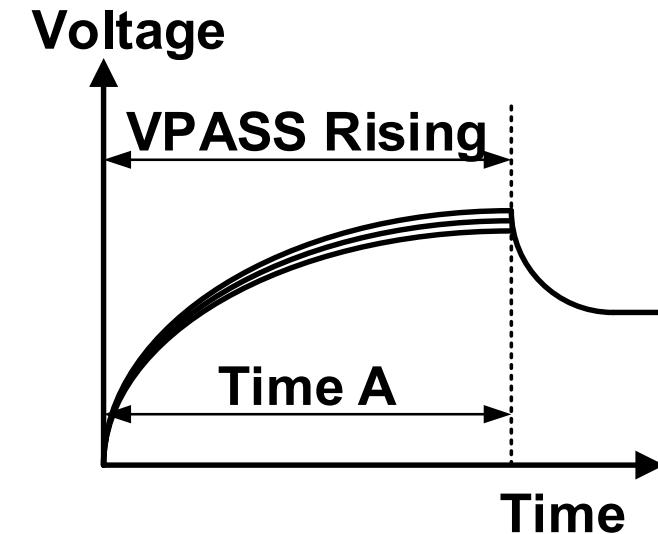
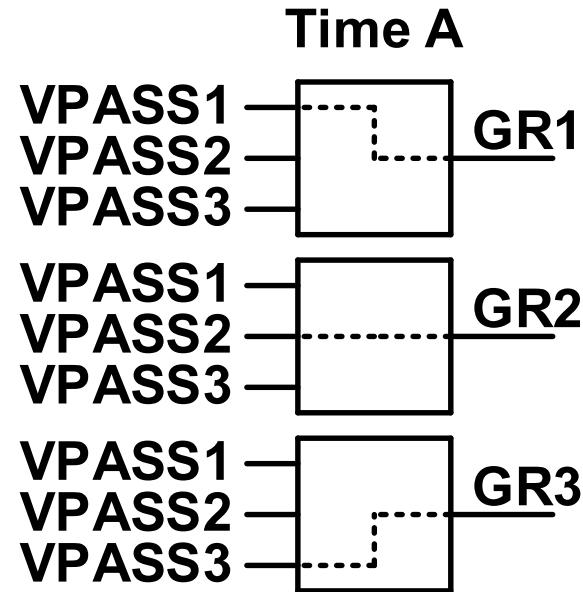
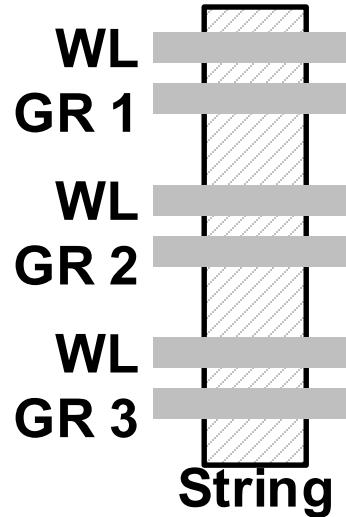
■ New Technologies

- Triple-verify Program (TPGM)
- Adaptive Unselected String Pre-charge (AUSP)
- Programmed Dummy String (PDS)
- **All-pass Rising (APR)**
- Plane-level Read Retry (PLRR)

■ Key Features

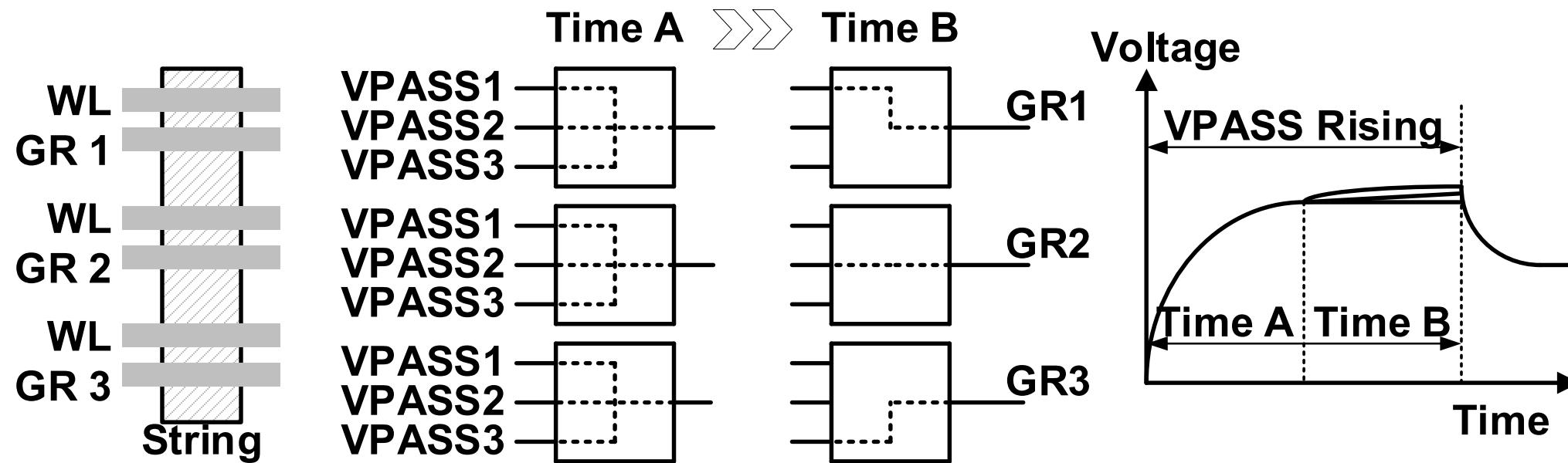
■ Conclusion

Conventional WL Rising



- Each cell needs different V_{PASS} due to different characteristics
- WLs are grouped to apply different VPASS levels
- One V_{PASS} source is selected and applied to WL groups

All-pass Rising (APR)



- Time A : all V_{PASS} sources are connected to reduce rising time
- Time B : one target V_{PASS} source is applied (conventional)
- The APR scheme reduces tR by around 2%

Outline

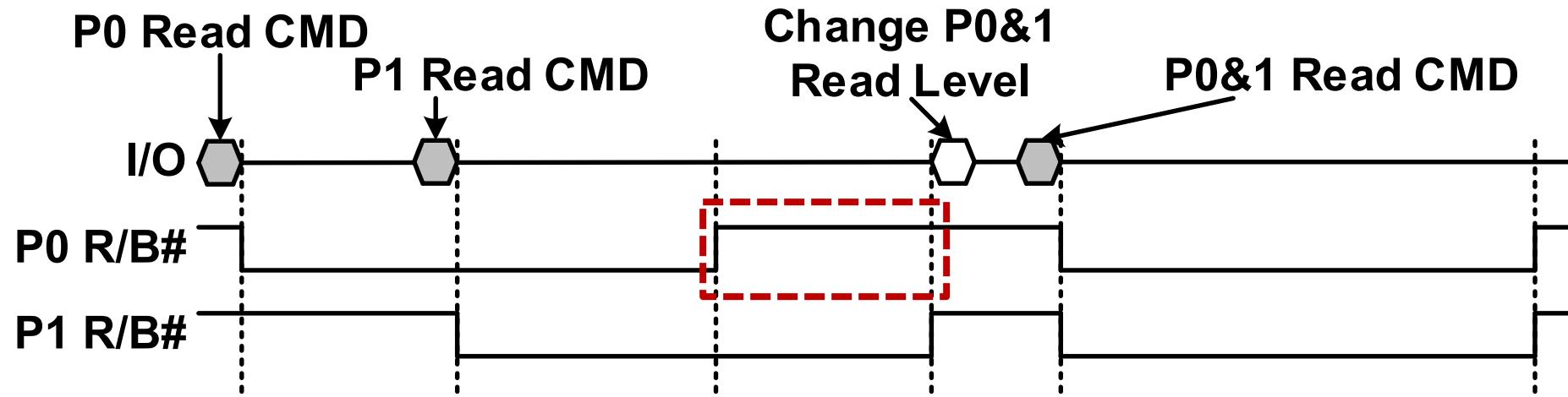
■ New Technologies

- Triple-verify Program (TPGM)
- Adaptive Unselected String Pre-charge (AUSP)
- Programmed Dummy String (PDS)
- All-pass Rising (APR)
- **Plane-level Read Retry (PLRR)**

■ Key Features

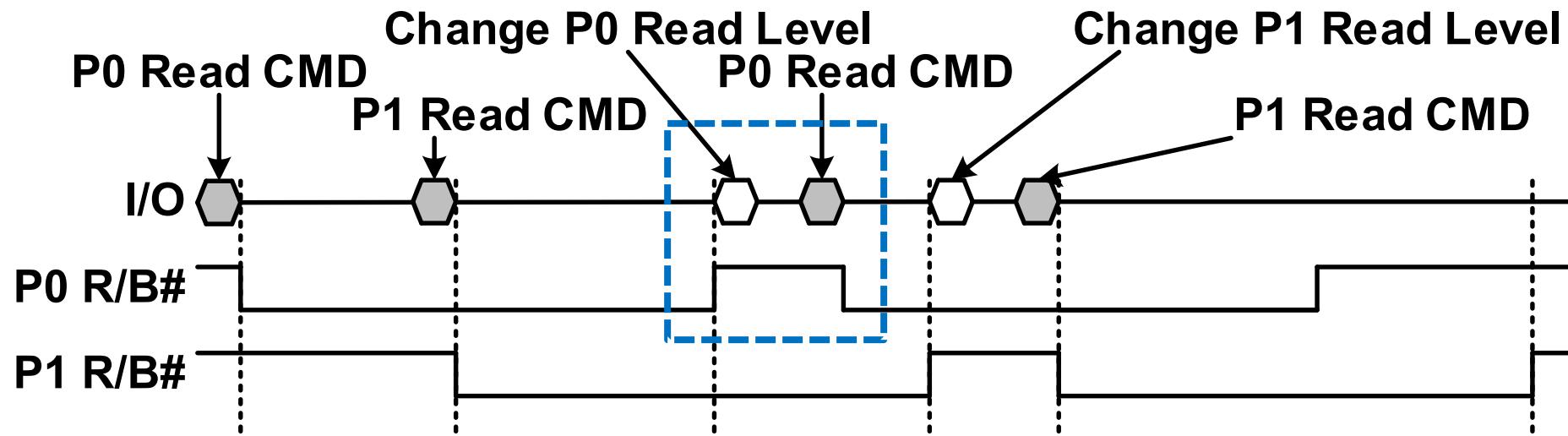
■ Conclusion

Conventional Read Retry



- Read Retry : read again with different read level
- Read level can be changed after completing all planes read
 - Plane0(P0) should wait until Plane1(P1) finishes read operation
 - the read performance is determined by the last plane terminated

Plane-level Read Retry (PLRR)



- Read level is changed regardless of other plane operations
- Compares to conventional RR, read performance is improved
 - Subsequent read commands can be issued immediately

Outline

■ New Technologies

- Triple-verify Program (TPGM)
- Adaptive Unselected String Pre-charge (AUSP)
- Programmed Dummy String (PDS)
- All-pass Rising (APR)
- Plane-level Read Retry (PLRR)

■ Key Features

■ Conclusion

Key Features



Die Photograph

	ISSCC 2022 [4]	This Work
# Bit/Cell	3	3
Capacity (Gb)	1024	1024
# of Planes	4	4
Page Size (KB/Page)	16	16
PGM Throughput (MB/s)	164	194
16KB tR (us)	45	34
IO Speed (Gbps)	2.4	2.4
Vccq (V)	1.2	1.2
Bit Density (Gb/mm ²)	11.55	>20

Feature Summary

Outline

■ New Technologies

- Triple-verify Program (TPGM)
- Adaptive Unselected String Pre-charge (AUSP)
- Programmed Dummy String (PDS)
- All-pass Rising (APR)
- Plane-level Read Retry (PLRR)

■ Key Features

■ Conclusion

Conclusion

- **A 1-Tb 3b/cell 3D-NAND Flash with 194MB/s write throughput for over 300 layers**
- **Introducing five new schemes for high-performance**
 - Triple-verify program (TPGM) is a scheme for better Cell V_{TH} dist. and tPROG
 - Adaptive unselected string pre-charge (AUSP) reduces disturb and enhances Cell V_{TH} dist. and tPROG
 - Programmed dummy string (PDS) reduces channel capacitance and reduces WL settling time for both tPROG and tR
 - All-pass rising (APR) technique reduces WL rising time and tR
 - Plane-level read retry (PLRR) improves Read Retry speed and enhances the Quality of Service

A 4-nm 16-Gb/s/pin Single-Ended PAM4 Parallel Transceiver with Switching-Jitter Compensation and Transmitter Optimization

**Jahoon Jin, Soo-Min Lee, Kyunghwan Min, Sodam Ju, Jihoon Lim,
Hyunsu Chae, Kwonwoo Kang, Yunji Hong, Yeongcheol Jeong,
Sang-Ho Kim, Jongwoo Lee, Joonsuk Kim**



Outline

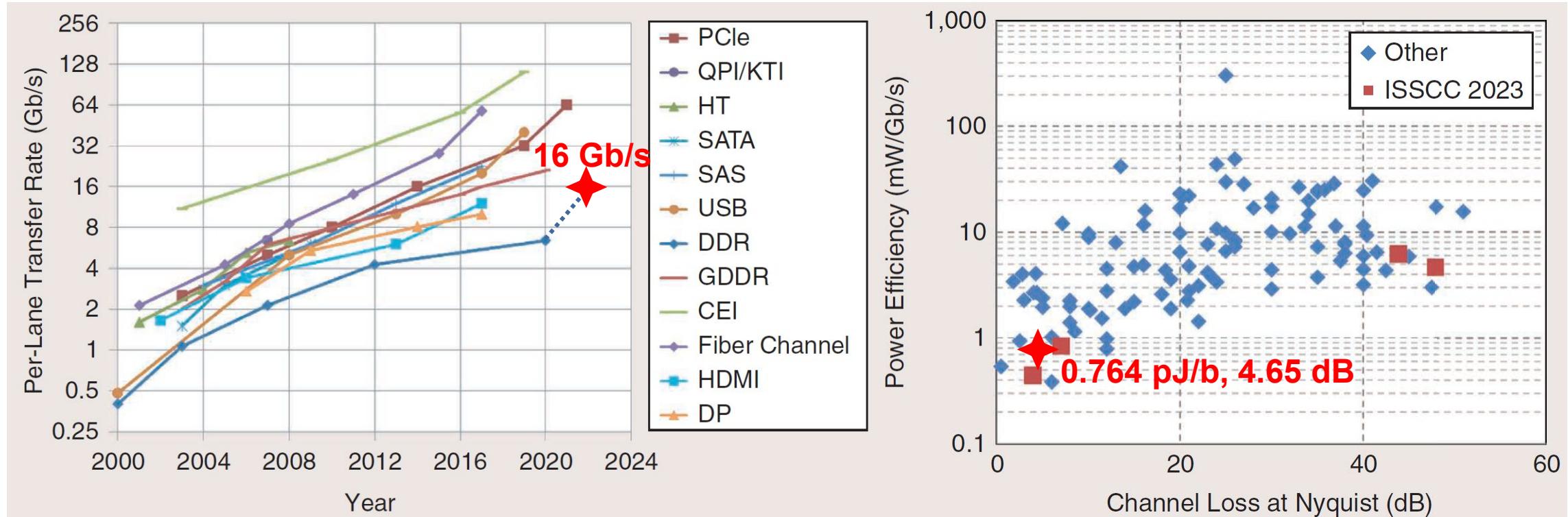
- **Introduction**
- **Concept of Switching-Jitter Compensation**
- **Design Optimization**
 - Fractionally-Spaced Feedforward Equalization
 - Relaxed Termination
- **Overall Architecture**
 - TX Implementation
 - RX Implementation
- **Measurements**
- **Conclusion**

Outline

- **Introduction**
- **Concept of Switching-Jitter Compensation**
- **Design Optimization**
 - Fractionally-Spaced Feedforward Equalization
 - Relaxed Termination
- **Overall Architecture**
 - TX Implementation
 - RX Implementation
- **Measurements**
- **Conclusion**

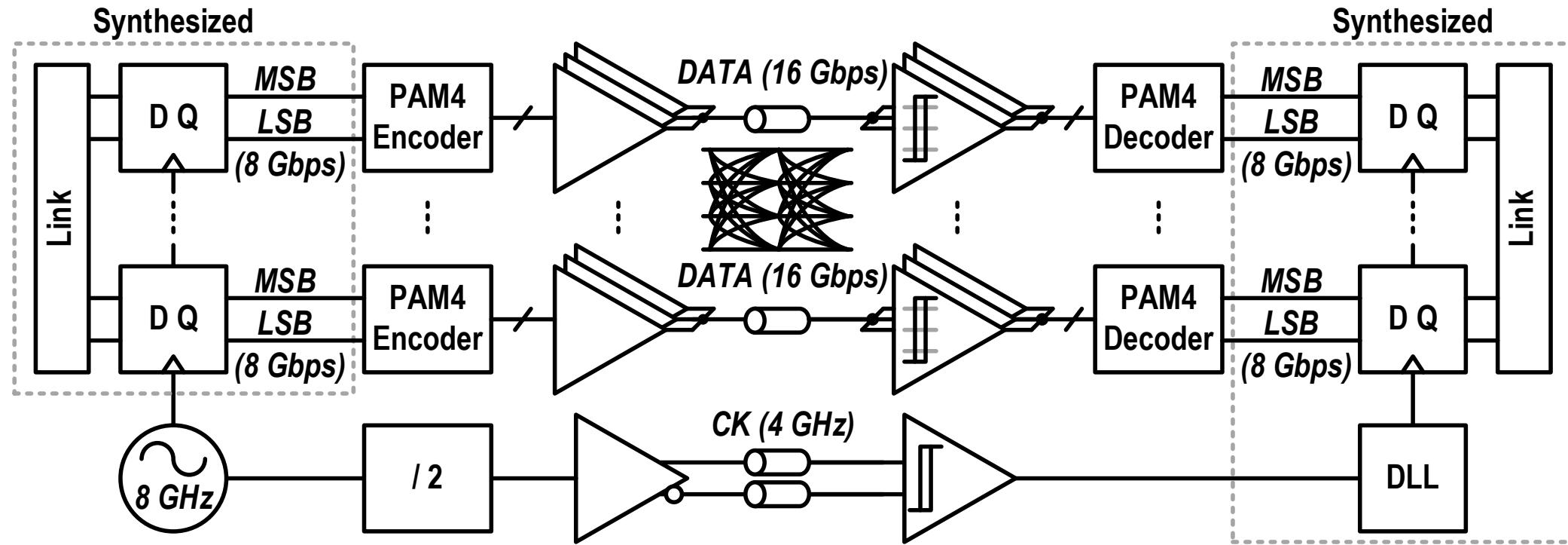
Trend

[IEEE Solid-State Circuits Magazine, 2023]

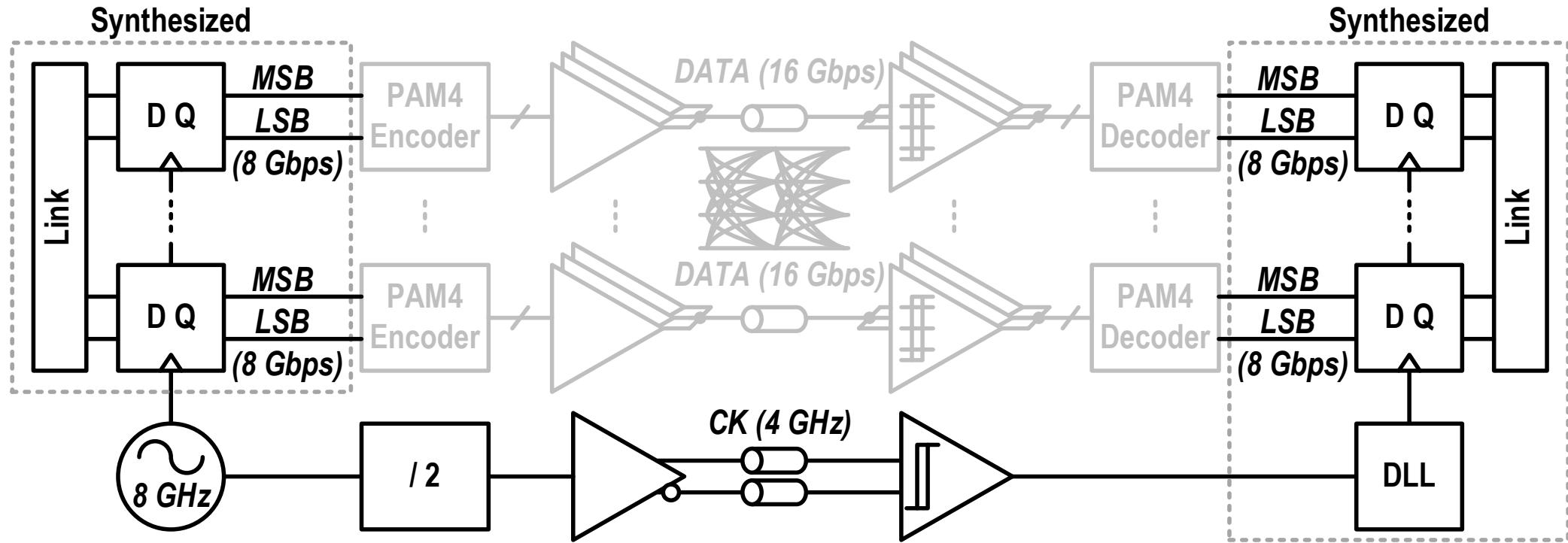


- Short-reach DDR-like serial interface
- This work achieves post LPDDR data rate.

Application: Short-Reach Parallel I/Os



Application: Short-Reach Parallel I/Os



Clock forwarded system

☺ **Fast data reconstruction**

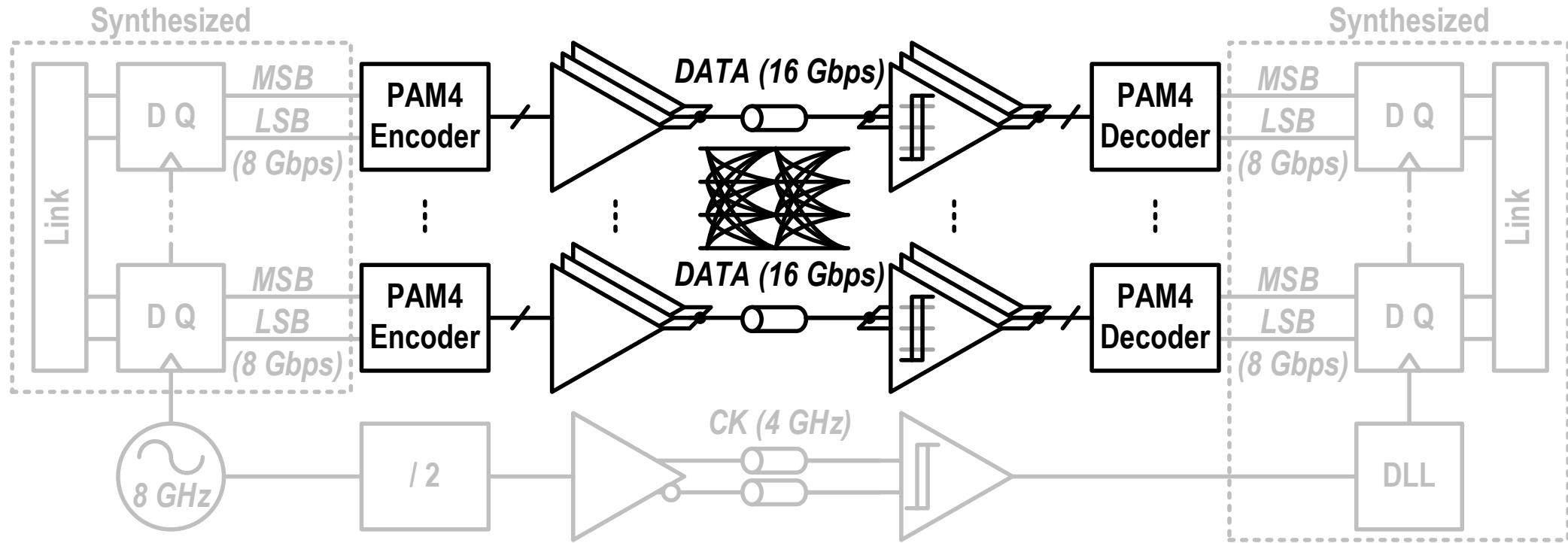
Analog asynchronous I/O design

☺ Power/area-efficient

PAM4

☺ Bandwidth extension

Application: Short-Reach Parallel I/Os



Clock forwarded system

☺ Fast data reconstruction

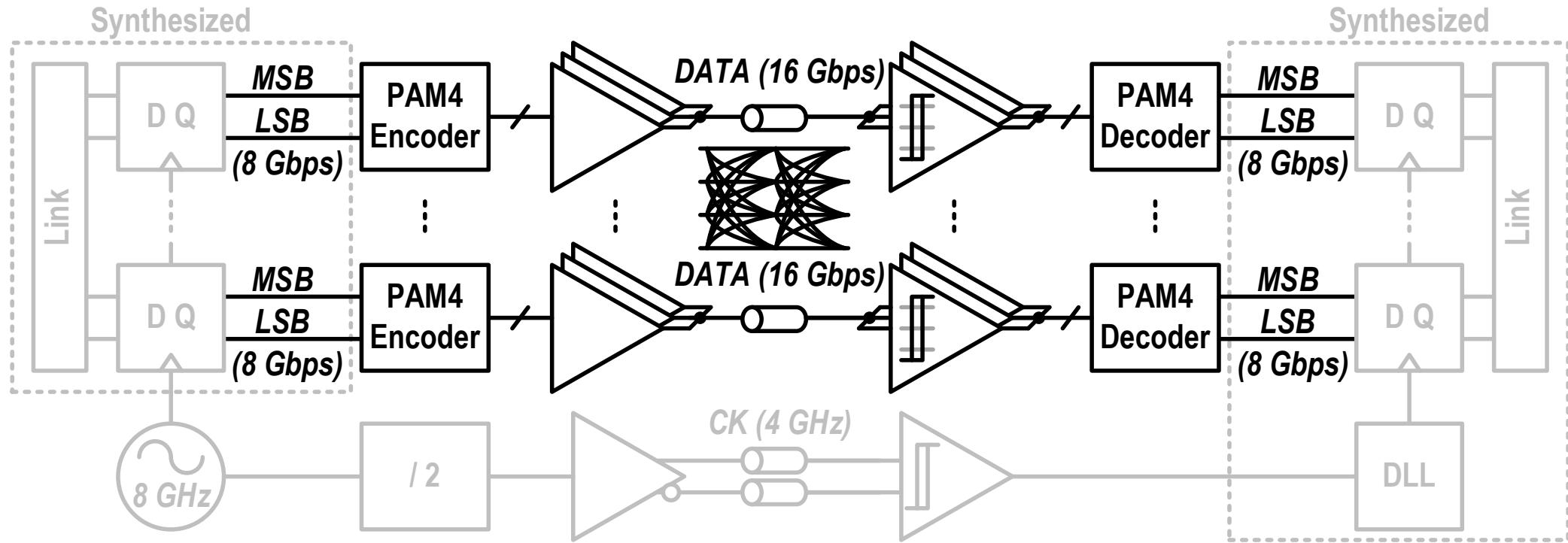
Analog asynchronous I/O design

☺ **Power/area-efficient**

PAM4

☺ Bandwidth extension

Application: Short-Reach Parallel I/Os



Clock forwarded system

☺ Fast data reconstruction

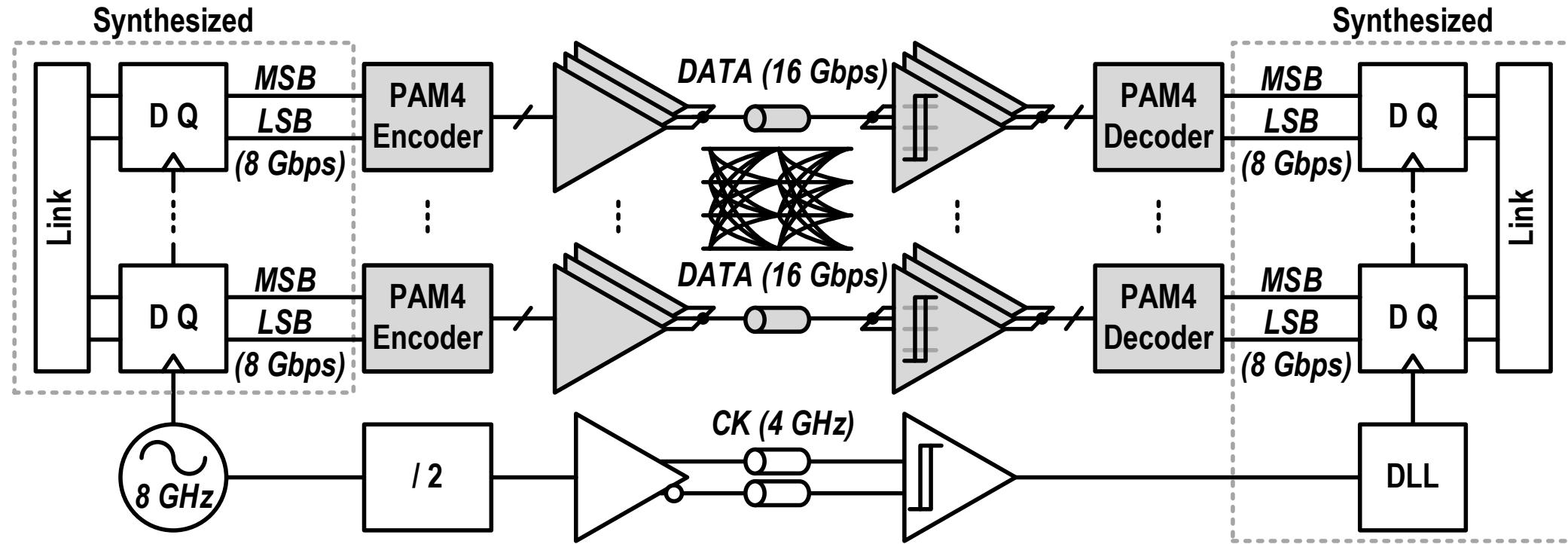
Analog asynchronous I/O design

☺ Power/area-efficient

PAM4

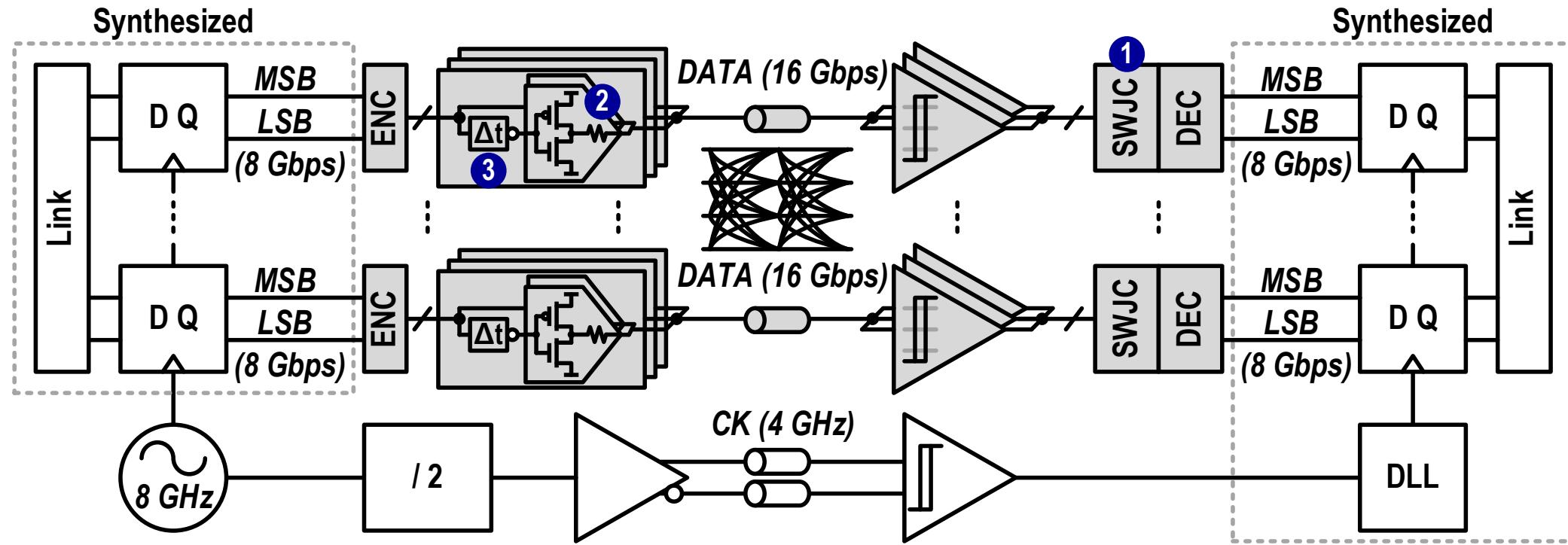
☺ **Bandwidth extension**

Focus: PAM4 I/O Design



PAM4	Solution	Goal
<ul style="list-style-type: none">😊 Throughput x2😢 SNR degradation😢 Switching jitter	<ol style="list-style-type: none">1. Switching jitter compensation2. Relaxed termination3. Fractionally-spaced FFE	<ul style="list-style-type: none">😊 Voltage margin ↑😊 Timing margin ↑

Focus: PAM4 I/O Design



PAM4	Solution	Goal
<ul style="list-style-type: none">😊 Throughput x2😢 SNR degradation😢 Switching jitter	<ol style="list-style-type: none">1. Switching jitter compensation2. Relaxed termination3. Fractionally-spaced FFE	<ul style="list-style-type: none">😊 Voltage margin ↑😊 Timing margin ↑

Outline

■ Introduction

■ Concept of Switching-Jitter Compensation

■ Design Optimization

- Fractionally-Spaced Feedforward Equalization
- Relaxed Termination

■ Overall Architecture

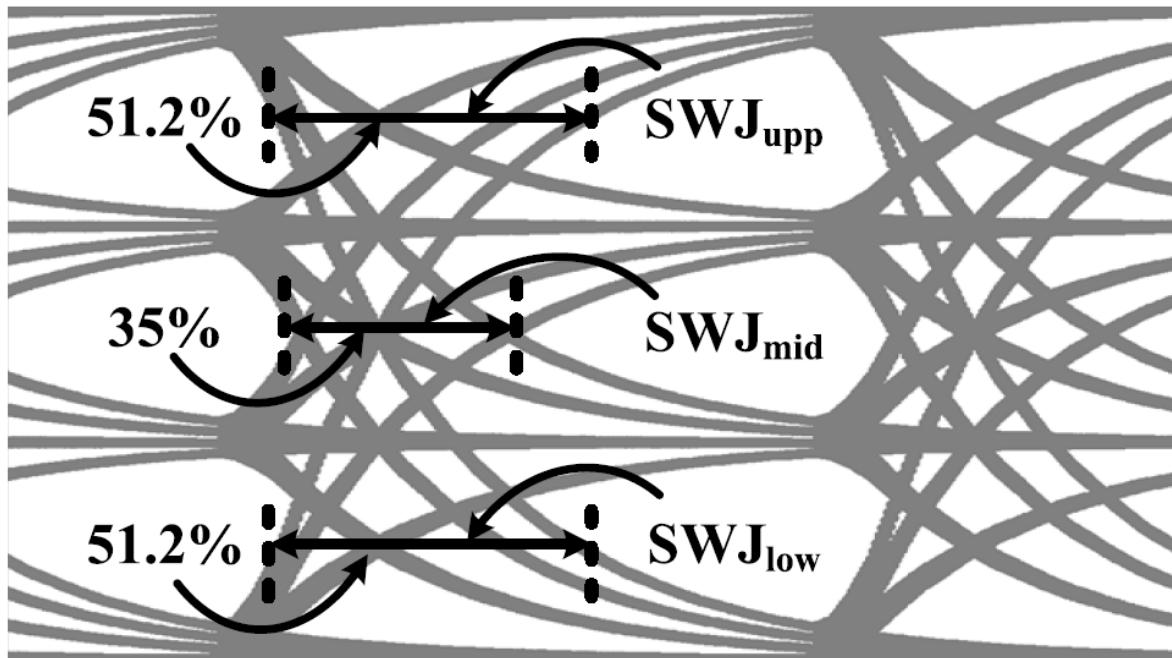
- TX Implementation
- RX Implementation

■ Measurements

■ Conclusion

Switching Jitters (SWJ) in PAM4 Signaling

1st-order low-pass filter with a Nyquist freq. cutoff

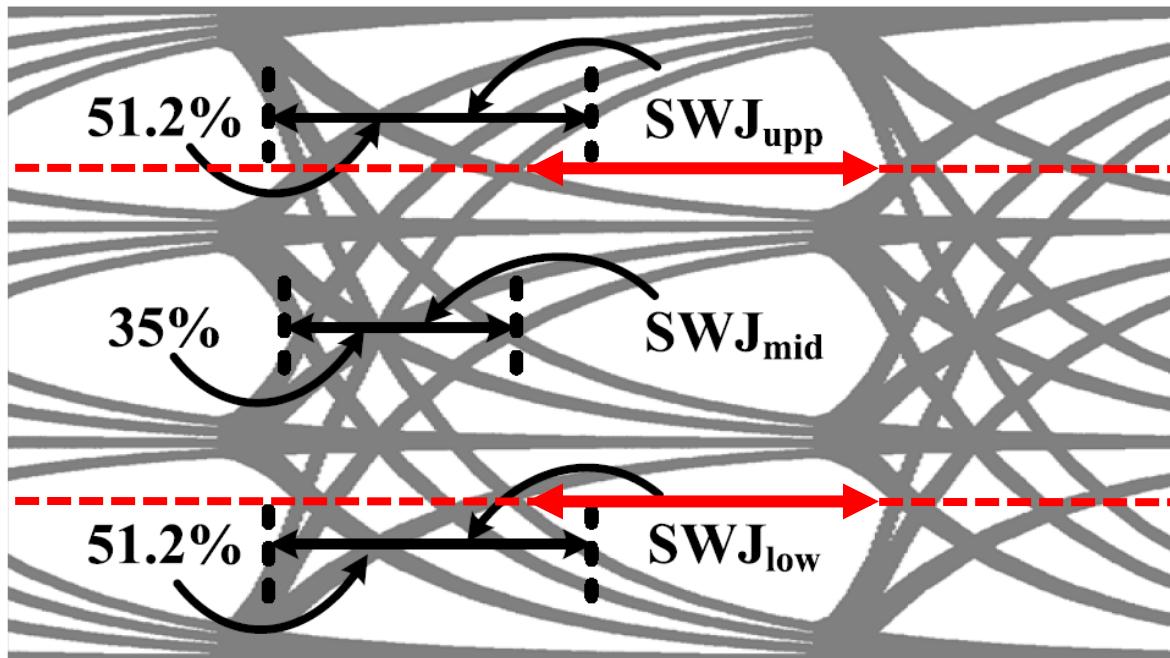


[X. Zheng, JSSC'20]

- Eyes are **asymmetrical** in shape.
- Upper/lower SWJ exceeds 50% of symbol interval.
- Mid SWJ reaches 35% of symbol interval.

Switching Jitters (SWJ) in PAM4 Signaling

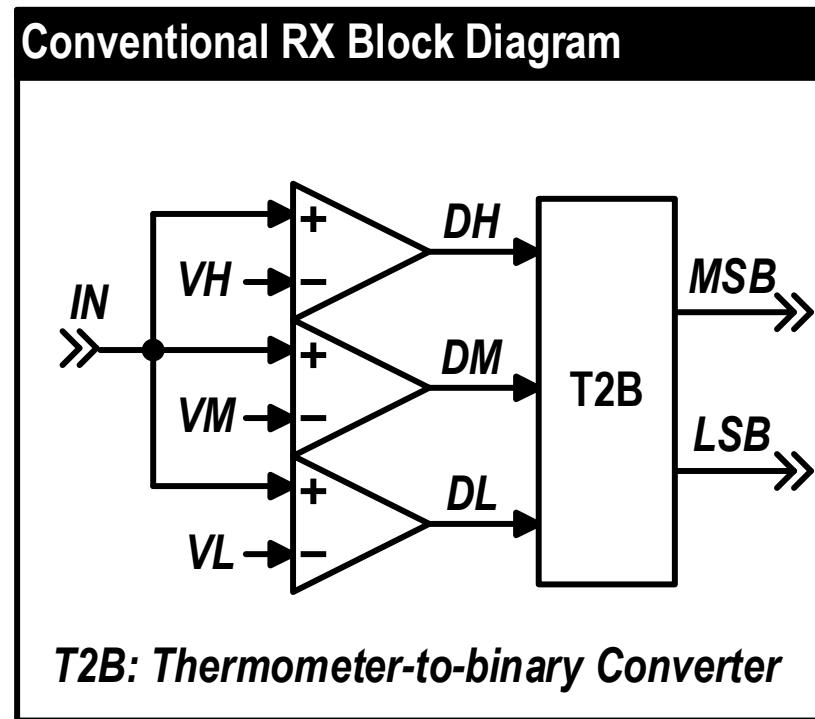
1st-order low-pass filter with a Nyquist freq. cutoff



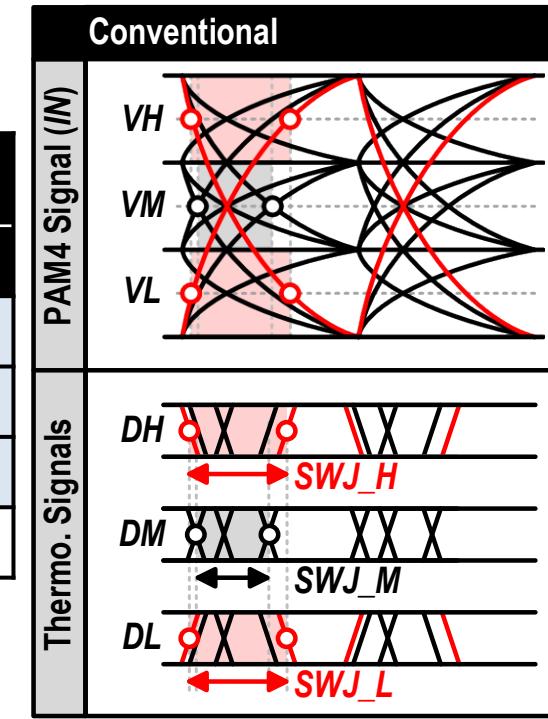
[X. Zheng, JSSC'20]

- Using **different VREFs** can be an option to increase timing margin but at the cost of **V margin sacrifice**.

Conventional Receiver

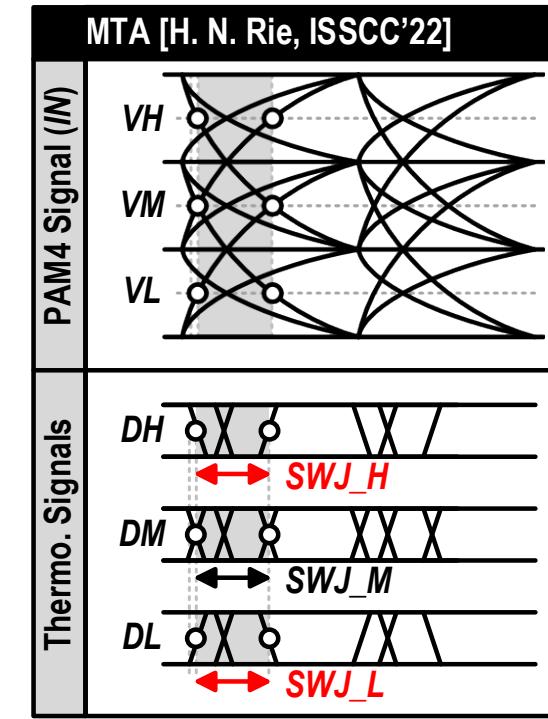
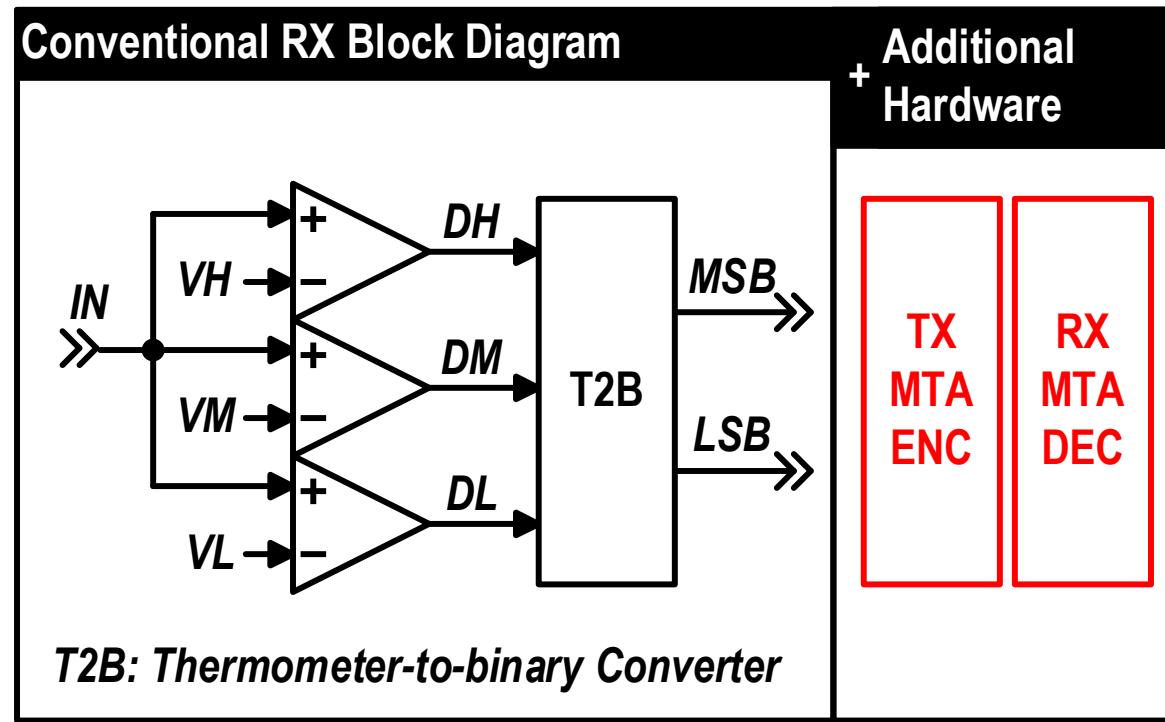


Voltage Level	Thermometer Code Conversion		
	DH	DM	DL
3	1	1	1
2	0	1	1
1	0	0	1
0	0	0	0



- Maximum transitions between the lowest and highest levels
- Contribute the most to SWJ.

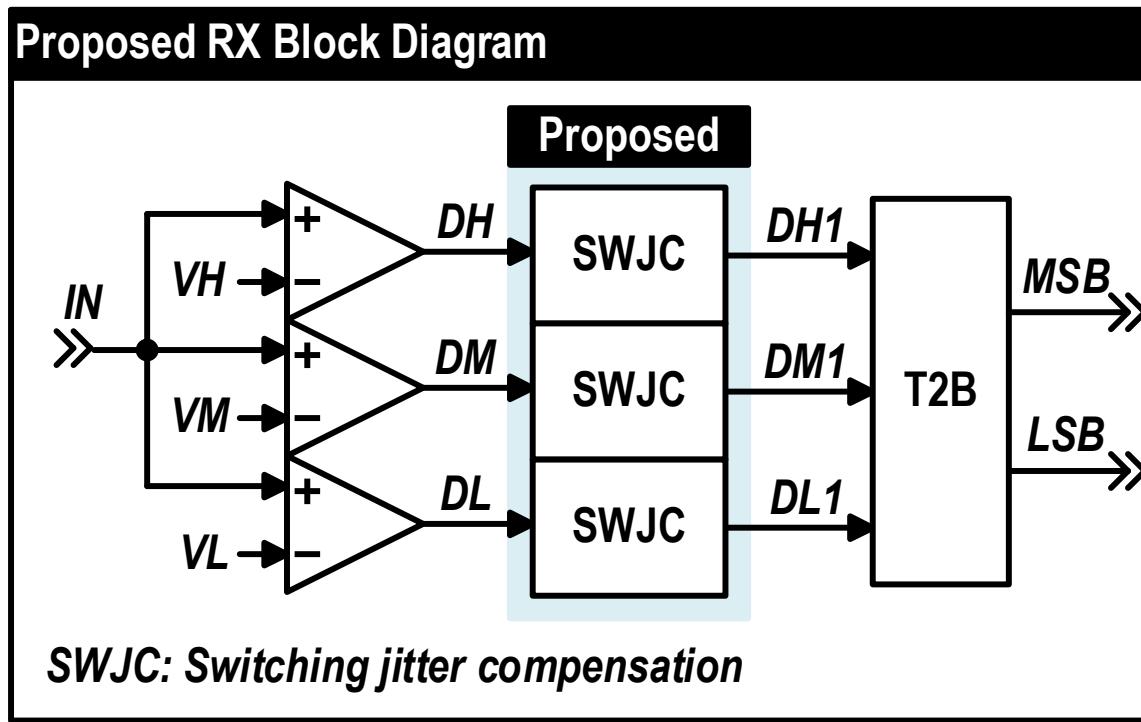
Maximum Transition Avoidance (MTA)



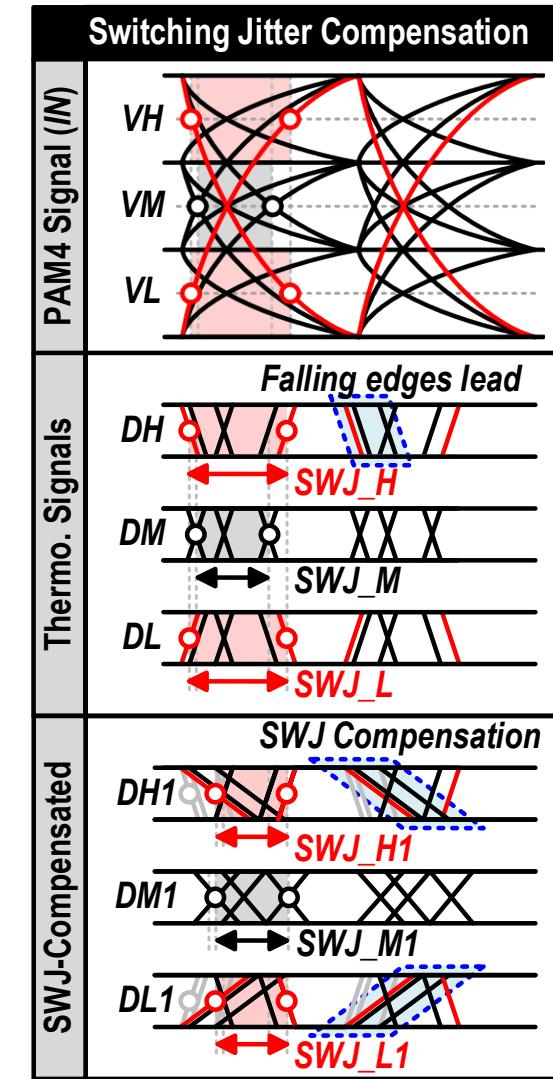
😊 Xtalk noise ↓
😊 SWJ ↓

😢 Additional encoder/decoder
😢 Pin efficiency ↓ (14 of 16 transitions are used.)

Switching Jitter Compensation



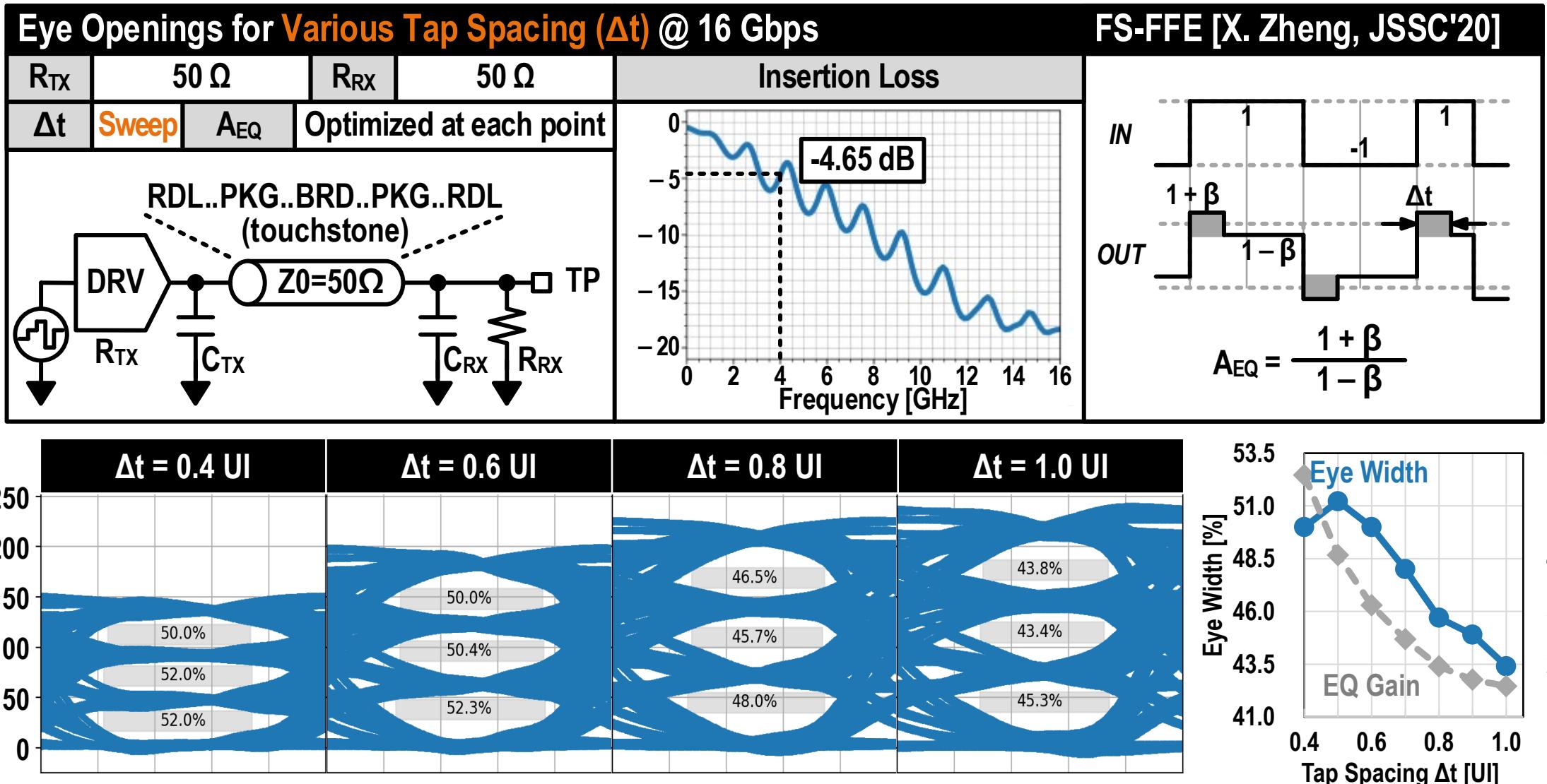
- 😊 Same Pin efficiency
- 😊 SWJ ↓ at the cost of small hardware



Outline

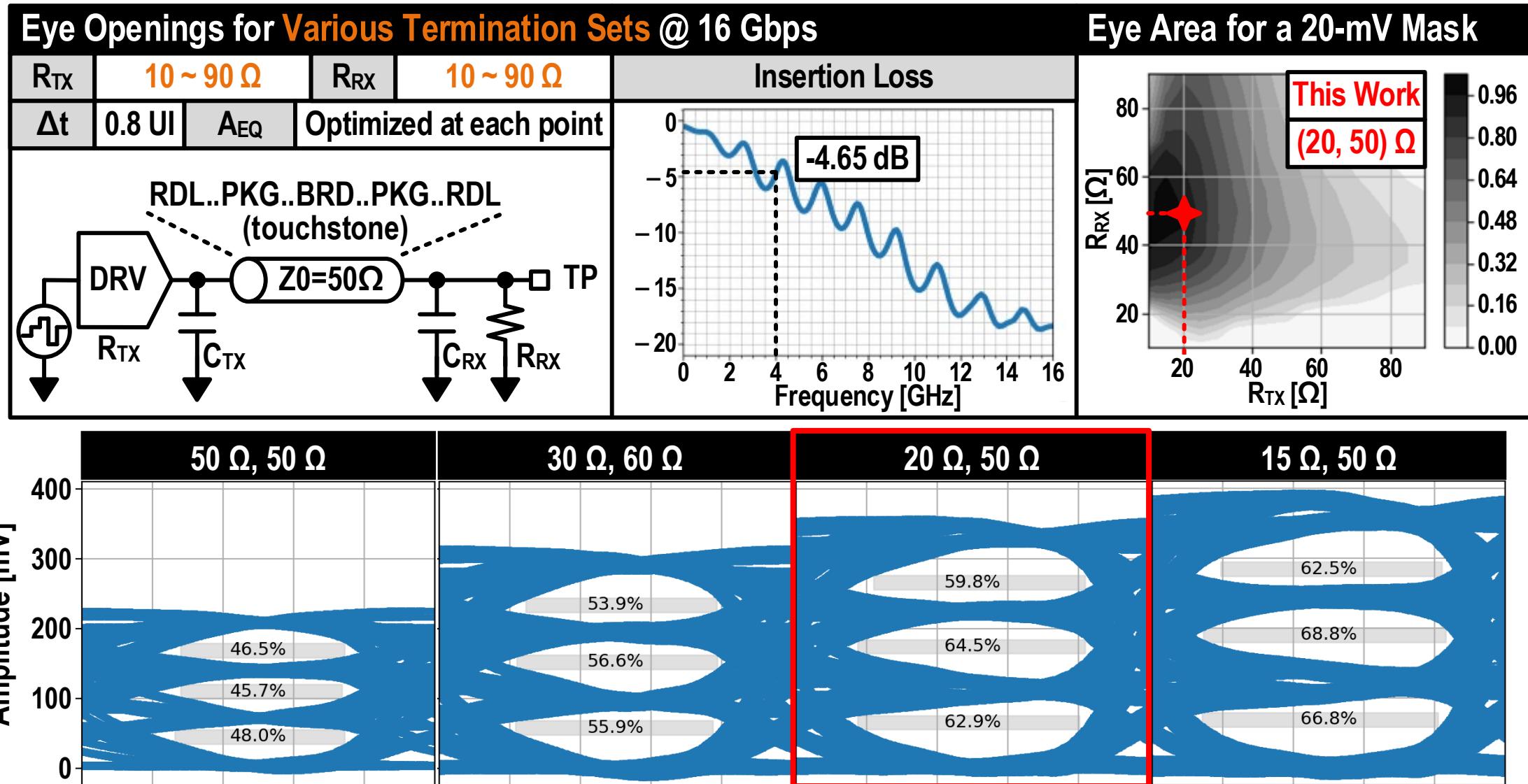
- **Introduction**
- **Concept of Switching-Jitter Compensation**
- **Design Optimization**
 - Fractionally-Spaced Feedforward Equalization
 - Relaxed Termination
- **Overall Architecture**
 - TX Implementation
 - RX Implementation
- **Measurements**
- **Conclusion**

Fractionally-Spaced FFE (FS-FFE)



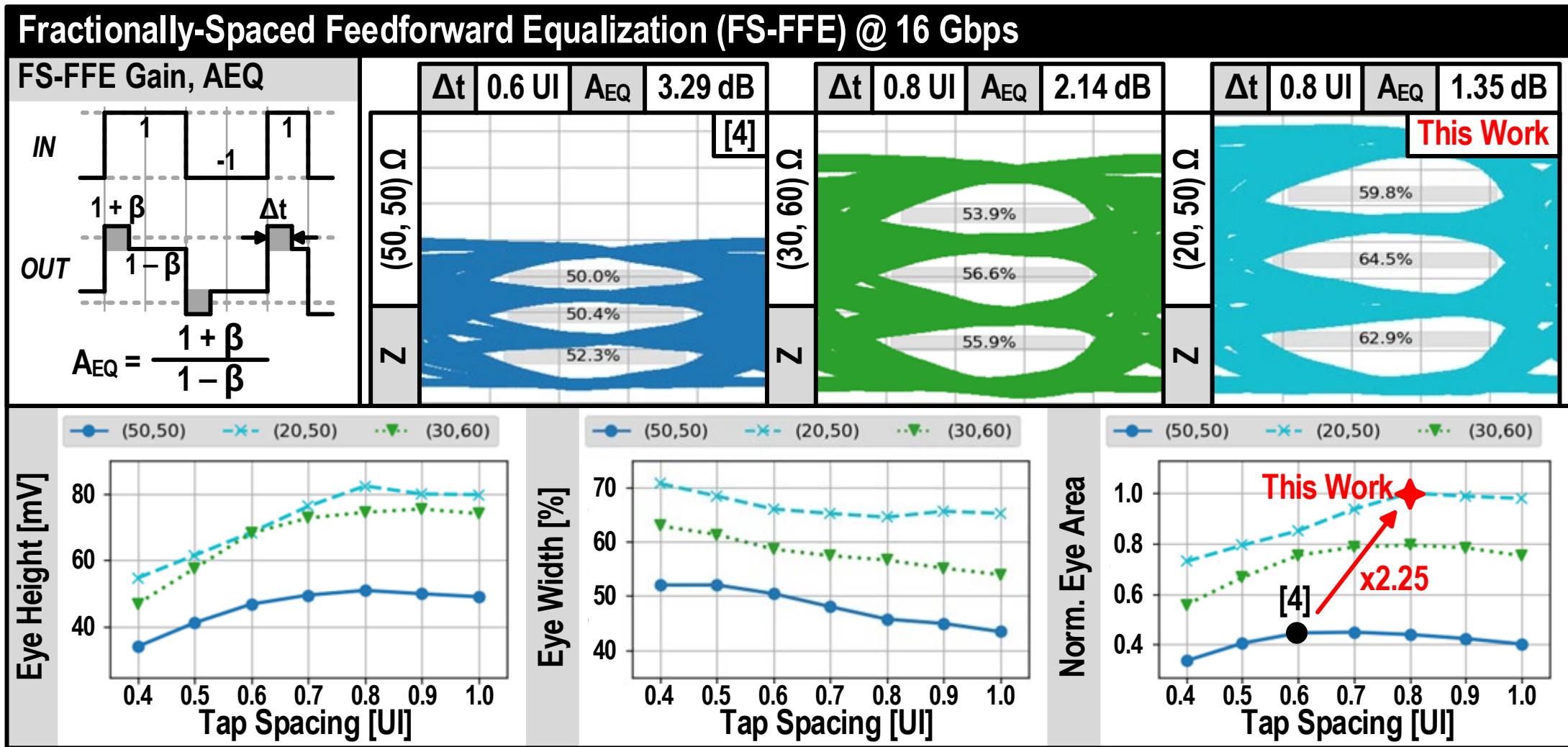
Relaxed Termination

[M. Choi, JSSC'18]



28.3: A 4-nm 16-Gb/s/pin Single-Ended PAM4 Parallel Transceiver with Switching-Jitter Compensation and Transmitter Optimization

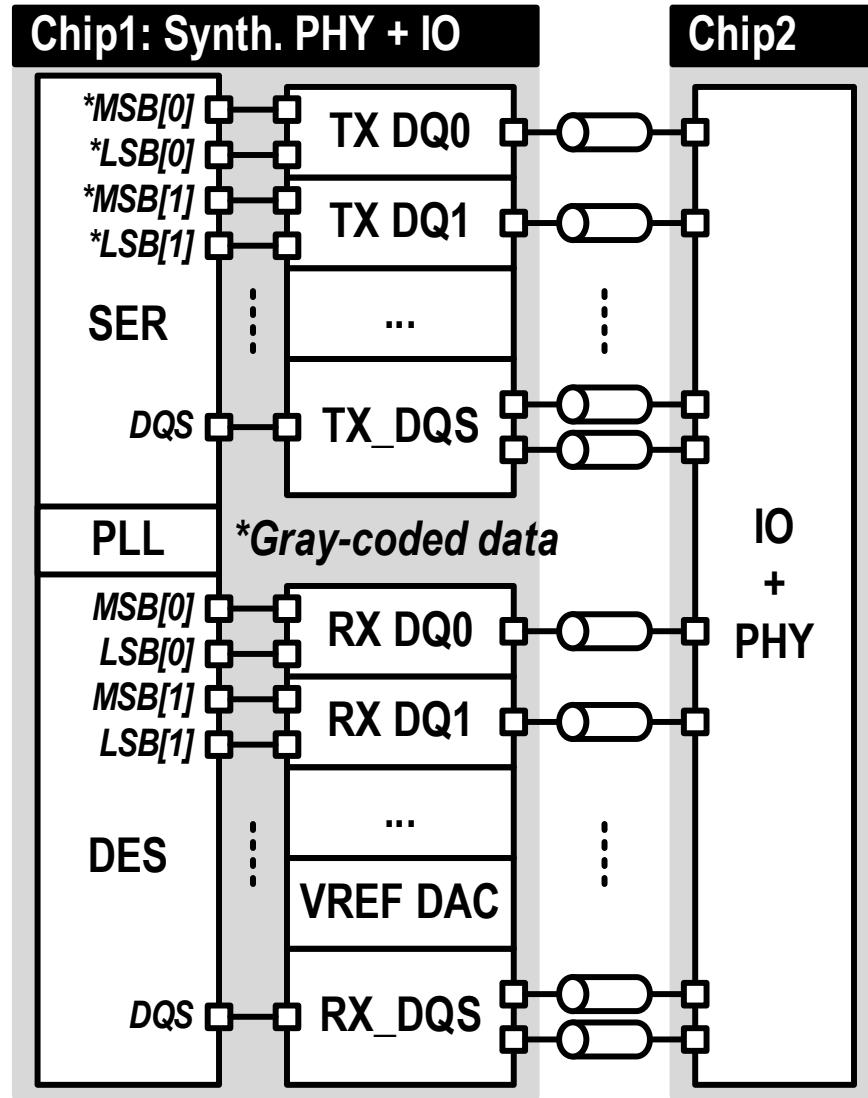
Optimization of FS-FFE Tap Spacing



Outline

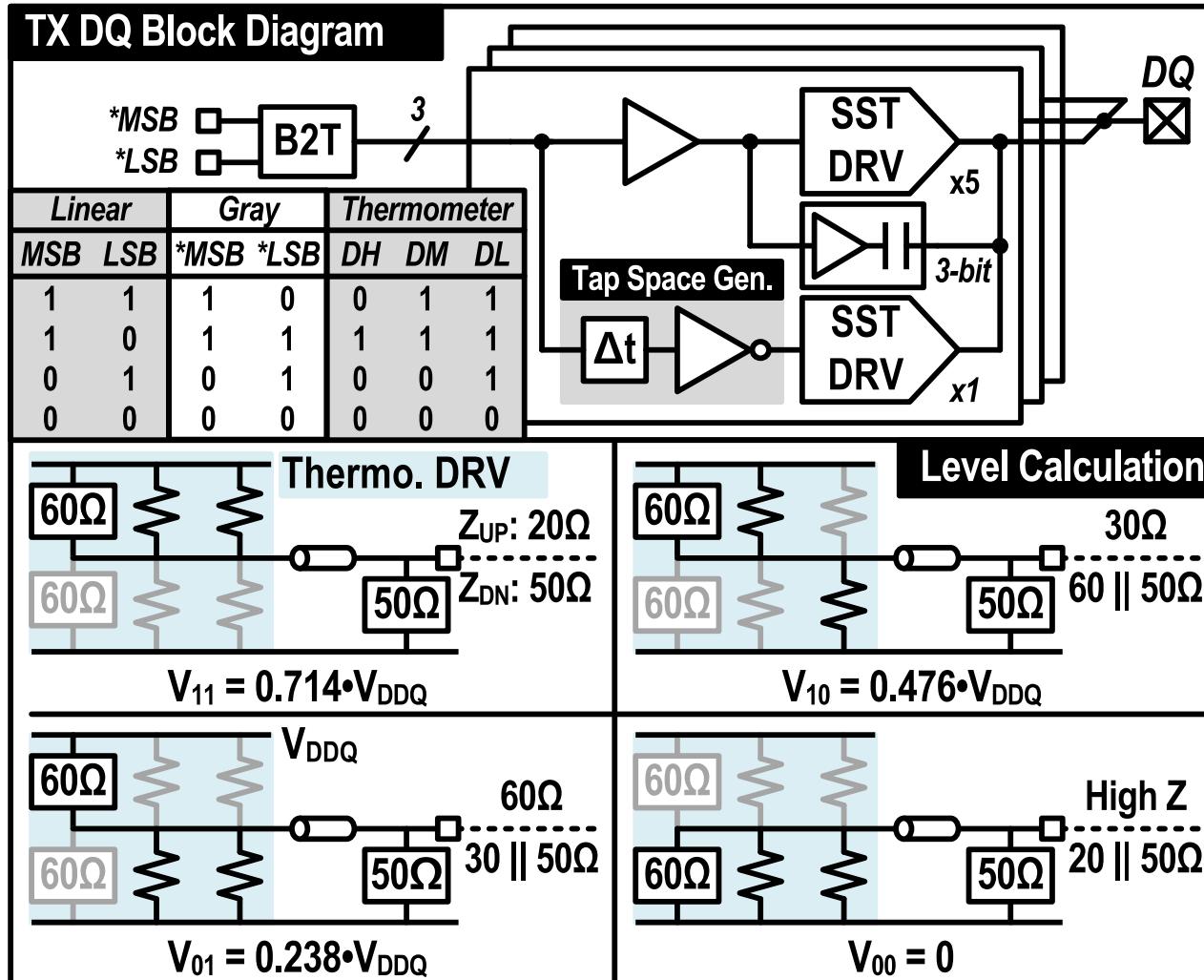
- **Introduction**
- **Concept of Switching-Jitter Compensation**
- **Design Optimization**
 - Fractionally-Spaced Feedforward Equalization
 - Relaxed Termination
- **Overall Architecture**
 - TX Implementation
 - RX Implementation
- **Measurements**
- **Conclusion**

Overall Architecture



- **Synthesized SER/DES**
 - 8-Gb/s MSB/LSB
- **16-Gb/s/pin throughput**
- **Gray coding**
 - RX simplicity

TX Architecture



Thermometer-code DRV

😊 High RLM

Relaxed Termination

😊 Eye opening ↑

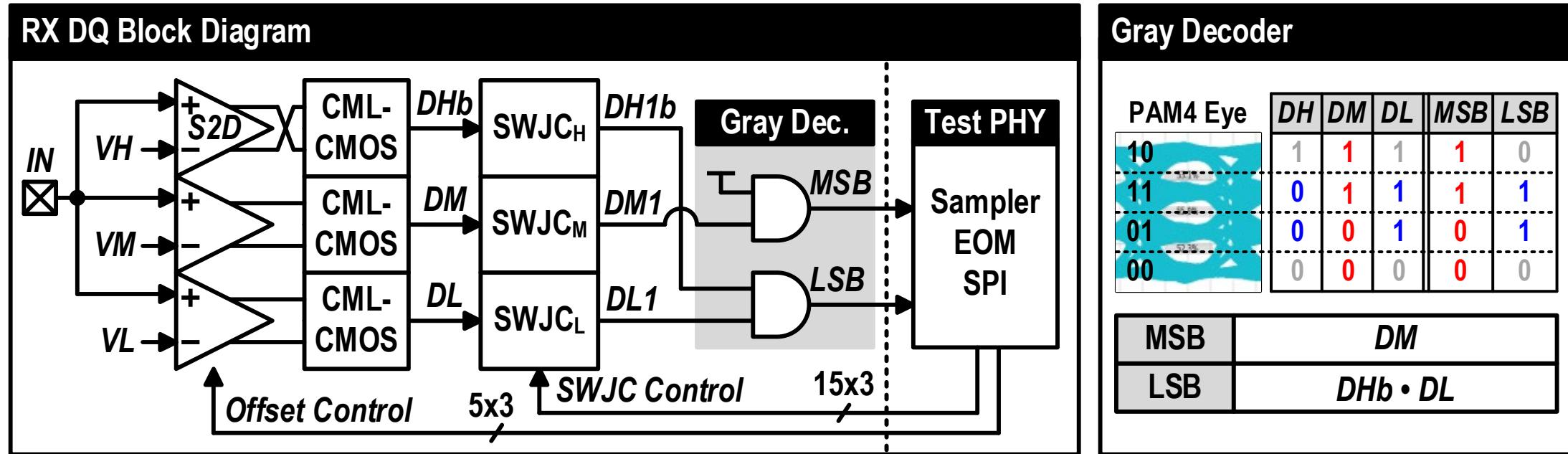
FS-FFE

😊 Bandwidth extension

Capacitive-peaking EQ

😊 Further strengthens high-freq. components

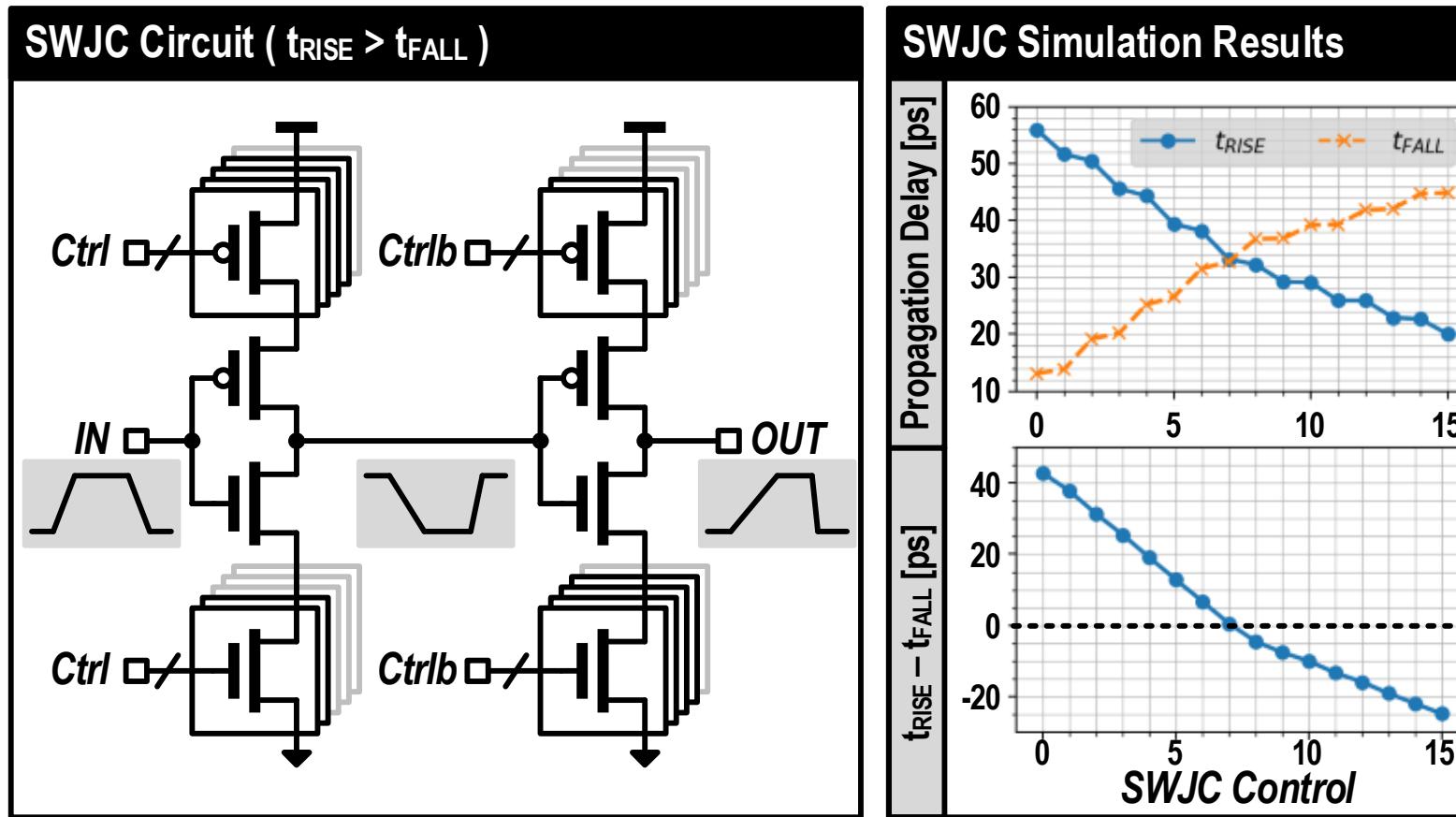
RX Architecture



■ MSB-LSB skew minimization

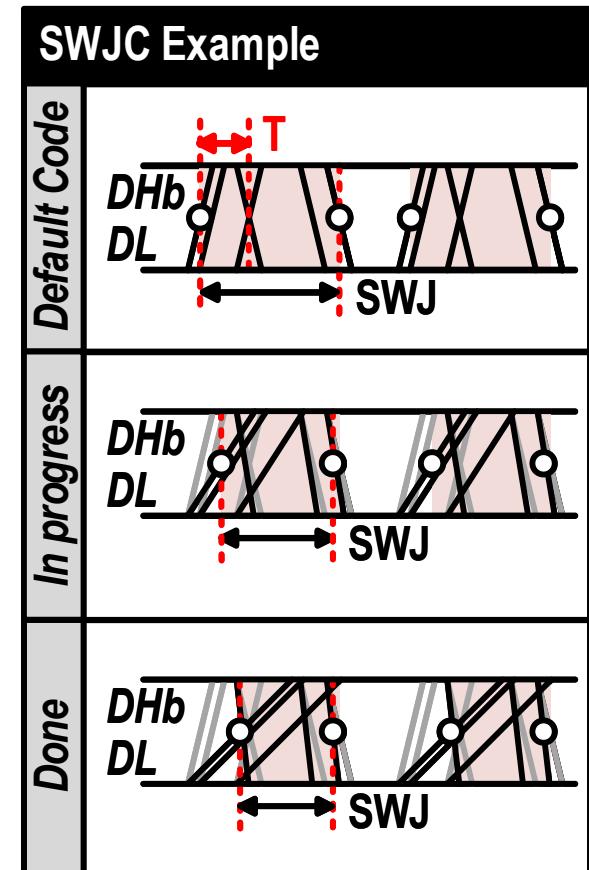
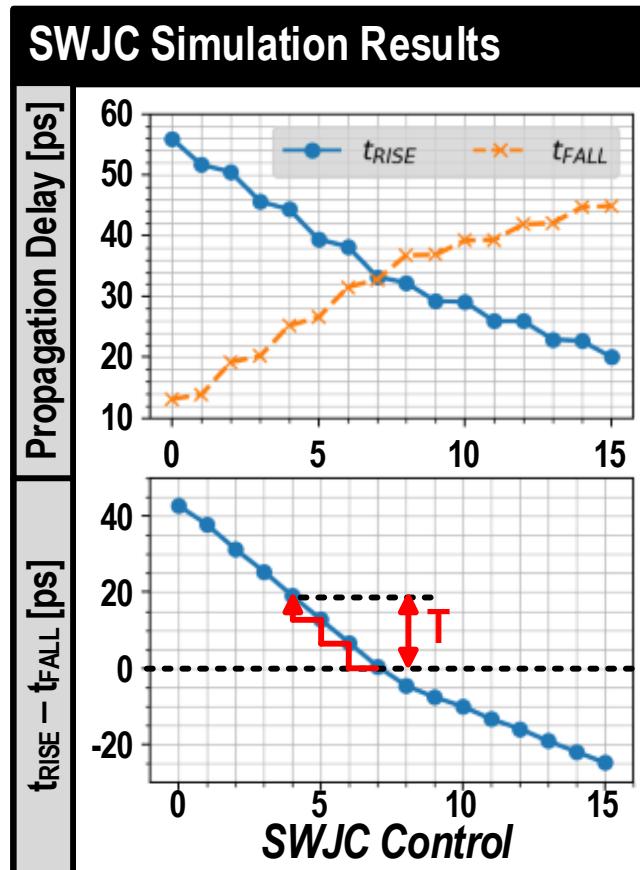
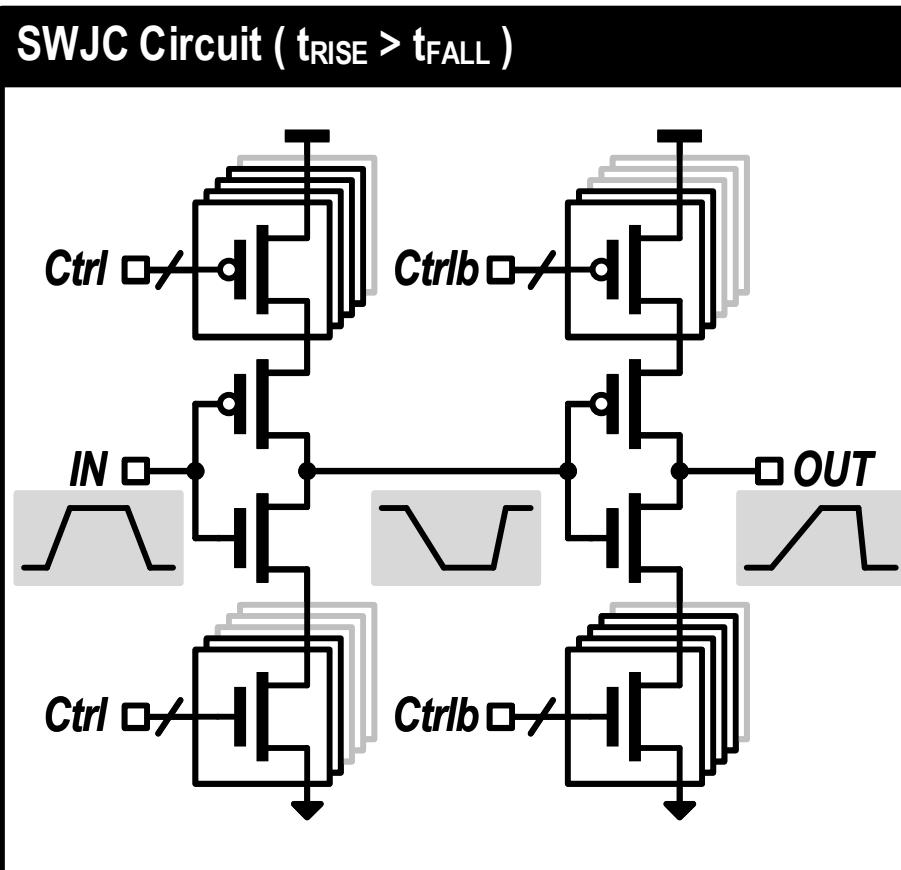
- Identical circuitry for MSB/LSB decoders
- Identical control of $SWJC_H$ and $SWJC_L$

SWJC Details



- Objective: Align the falling edges and rising edges.
- How: Adjust the propagation delay of transitions.

SWJC Details



- Objective: Align the falling edges and rising edges.
- How: Adjust the propagation delay of transitions.

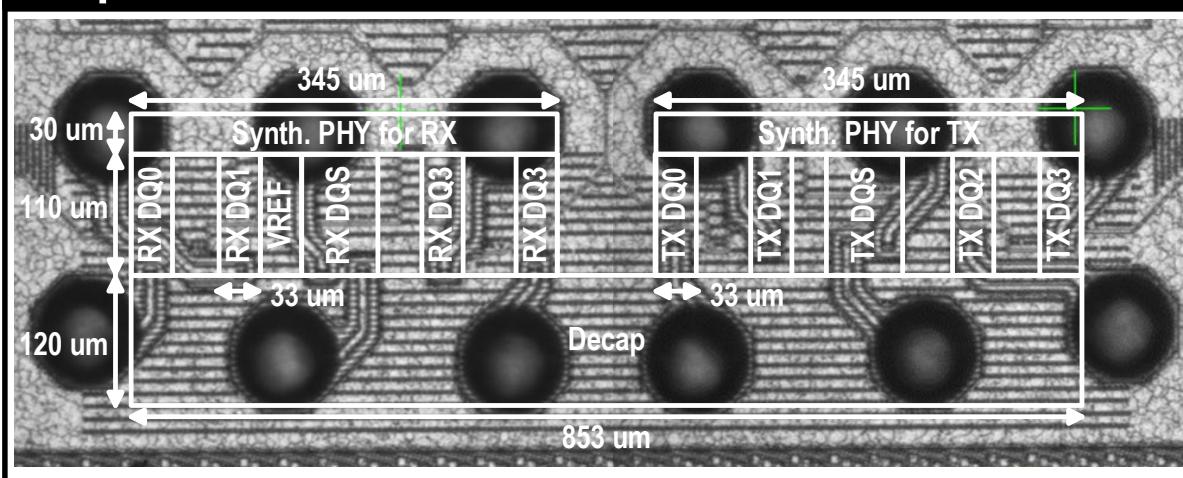
Outline

- **Introduction**
- **Concept of Switching-Jitter Compensation**
- **Design Optimization**
 - Fractionally-Spaced Feedforward Equalization
 - Relaxed Termination
- **Overall Architecture**
 - TX Implementation
 - RX Implementation
- **Measurements**
- **Conclusion**

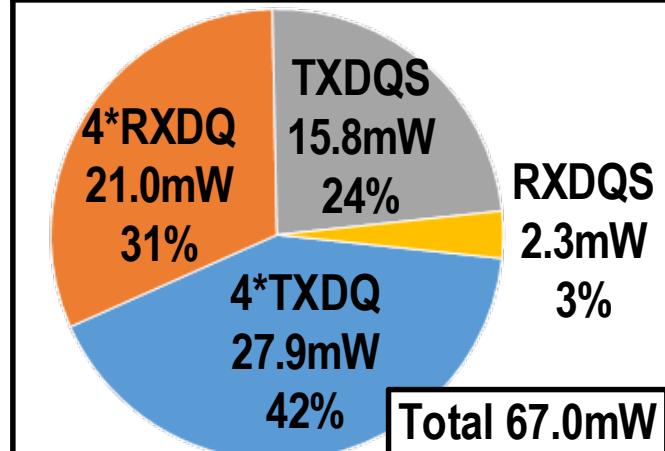
Measurement Results – Chip Info.

- Process: 4nm FinFET
- Prototype IC: 64 Gb/s
 - TX(4DQ + 1DQS)
 - RX(4DQ + 1DQS + VREF)
- Data rate: 16 Gb/s/pin
- Energy efficiency: 0.764 pJ/b (DQ only)

Chip Photo



Power Breakdown of 64 Gb/s Prototype



	FoM [pJ/b]
TXDQ	0.436
RXDQ	0.328
DQS	0.282
ALL	1.046

Measurement Results – TX

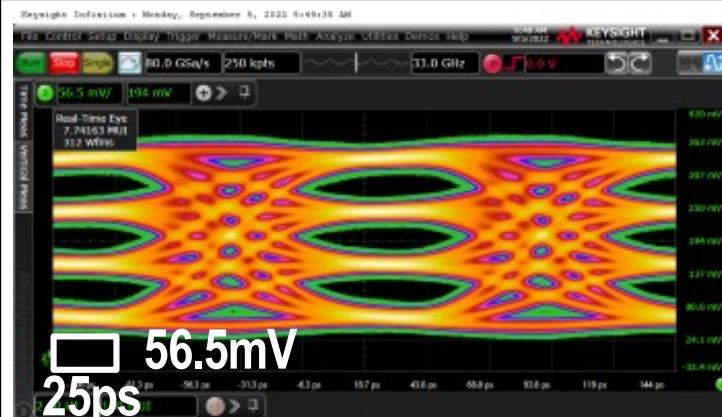
TX Eye Measurements

No EQ



Eye	Height [mV]	Width [ps]
Top	14.42	31.46
Middle	15.68	38.20
Bottom	14.43	31.74

+ FS-FFE



Eye	Height [mV]	Width [ps]
Top	40.88	50.28
Middle	39.68	58.15
Bottom	40.88	51.12

+ FS-FFE + C-peaking



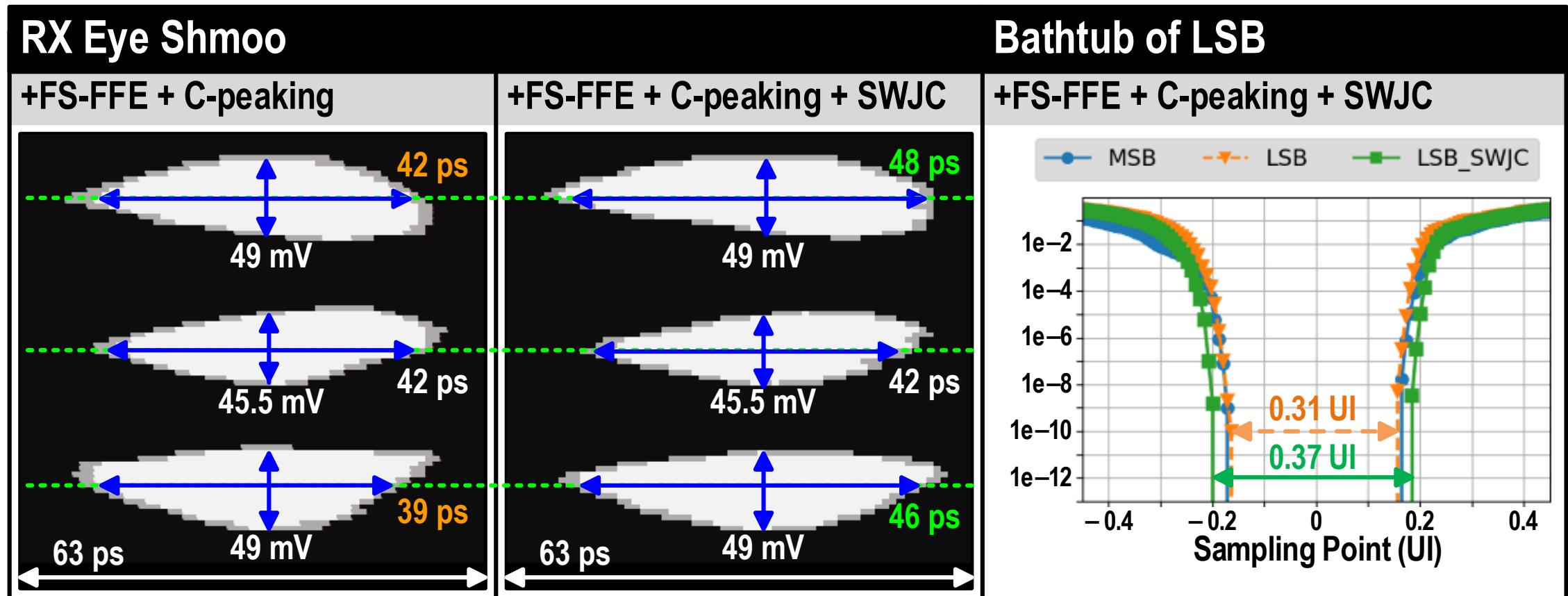
Eye	Height [mV]	Width [ps]
Top	40.87	62.93
Middle	39.67	67.14
Bottom	37.26	60.11

0.40 UI

+0.08UI

0.48 UI

Measurement Results – RX



- SWJC improves LSB opening from 0.31 UI to 0.37 UI.

Outline

- **Introduction**
- **Concept of Switching-Jitter Compensation**
- **Design Optimization**
 - Fractionally-Spaced Feedforward Equalization
 - Relaxed Termination
- **Overall Architecture**
 - TX Implementation
 - RX Implementation
- **Measurements**
- **Conclusion**

Conclusion

	This Work	ISSCC'21 25.3 [1]	ASSCC'21 17.2 [2]	VLSI'22 [3]	CICC'22 J. Kim
Process	4nm FinFET	1Ynm CMOS	28nm CMOS	28nm CMOS	28nm CMOS
Supply	0.75V / 0.6V	1.35V	1V	1.2V / 0.95V	1.2V / 1.2V
Data Rate	8G/16G	11G/22G	12G/24G	40G	60G
Modulation	NRZ/PAM4	NRZ/PAM4	NRZ/PAM4	PAM4	PAM4
Termination	TX20 - RX50	TX40 - RX40	N/A	N/A	1:1
Equalization	1T-FS-FFE C-peaking	1T-FFE CTLE	1T-DFE	T-Coil, 2T-FFE CTLE, 4T-DFE	2T-FFE
SWJ Reduction	SWJC	X	X	X	MTA
Eye Opening (BER)	0.37 UI (1e-12)	0.31 UI (1e-12)	0.23 UI ^(C) (N/A)	0.3 UI (1e-11)	0.2 UI (1e-6)
FoM [pJ/b]	0.764 ^(A) , 1.046 ^(B)	N/A	N/A	2.02	1.67 ^(D)

(A) DQ I/O only

(B) 64Gb/s prototype IC including DQS I/Os

(D) TX only

(C) Estimated from the write shmoo plot @ 24 Gb/s/pin

Conclusion

	This Work	ISSCC'21 25.3 [1]	ASSCC'21 17.2 [2]	VLSI'22 [3]	CICC'22 J. Kim
Process	4nm FinFET	1Vnm CMOS	28nm CMOS	28nm CMOS	28nm CMOS
Supply	0.75V / 0.6V				2V / 1.2V
Data Rate	8G/16G				60G
Modulation	NRZ/PAM4	NRZ/PAM4	NRZ/PAM4	PAM4	PAM4
Termination	TX20 - RX50	TX40 - RX40	N/A	N/A	1:1
Equalization	1T-FS-FFE C-peaking	1T-FFE CTLE	1T-DFE	T-Coil, 2T-FFE CTLE, 4T-DFE	2T-FFE
SWJ Reduction	SWJC	X	X	X	MTA
Eye Opening (BER)	0.37 UI (1e-12)	0.31 UI (1e-12)	0.23 UI ^(C) (N/A)	0.3 UI (1e-11)	0.2 UI (1e-6)
FoM [pJ/b]	0.764 ^(A) , 1.046 ^(B)	N/A	N/A	2.02	1.67 ^(D)

Achieves **2.25x wider eye opening.**

(A) DQ I/O only

(B) 64Gb/s prototype IC including DQS I/Os

(D) TX only

(C) Estimated from the write shmoo plot @ 24 Gb/s/pin

Conclusion

	This Work	ISSCC'21 25.3 [1]	ASSCC'21 17.2 [2]	VLSI'22 [3]	CICC'22 J. Kim
Process	4nm FinFET	1Ynm CMOS	28nm CMOS	28nm CMOS	28nm CMOS
Supply	0.75V / 0.6V	1.35V	1V	1.2V / 0.95V	1.2V / 1.2V
Data Rate	8G/16G	11G/22G	12G/24G	40G	60G
Modulation	NRZ/PAM4	NRZ/PAM4	NRZ/PAM4	PAM4	PAM4
Termination	TX20 - RX50	Improves timing margin by 0.06 UI while maintaining the same throughput.			
Equalization	1T-FS-FFE C-peaking	CTLE, 4T-DFE			
SWJ Reduction	SWJC	X	X	X	MTA
Eye Opening (BER)	0.37 UI (1e-12)	0.31 UI (1e-12)	0.23 UI ^(C) (N/A)	0.3 UI (1e-11)	0.2 UI (1e-6)
FoM [pJ/b]	0.764 ^(A) , 1.046 ^(B)		N/A	N/A	2.02
					1.67 ^(D)

(A) DQ I/O only

(B) 64Gb/s prototype IC including DQS I/Os

(D) TX only

(C) Estimated from the write shmoo plot @ 24 Gb/s/pin

A 4nm 1.15TB/s HBM3 Interface with Resistor-Tuned Offset-Calibration and In-Situ Margin-Detection

Kwanyeob Chae, Jiyeon Park, Jaegeun Song, Billy Koo, Jihun Oh,
Shinyoung Yi, Won Lee, Dongha Kim, Taekyung Yeo, Kyeongkeun Kang,
Sangsoo Park, Eunsu Kim, Sukhyun Jung, Sanghune Park, Sungcheol
Park, Mijung Noh, Hyogyuem Rhew, Jongshin Shin

Samsung Electronics, Korea

Outline

- Introduction
- Proposed Architecture
- Implementation
- Measurement Results
- Conclusion

Outline

■ Introduction

- HBM Interface Trend
- Function of Memory Interface
- Technical Challenges

■ Proposed Architecture

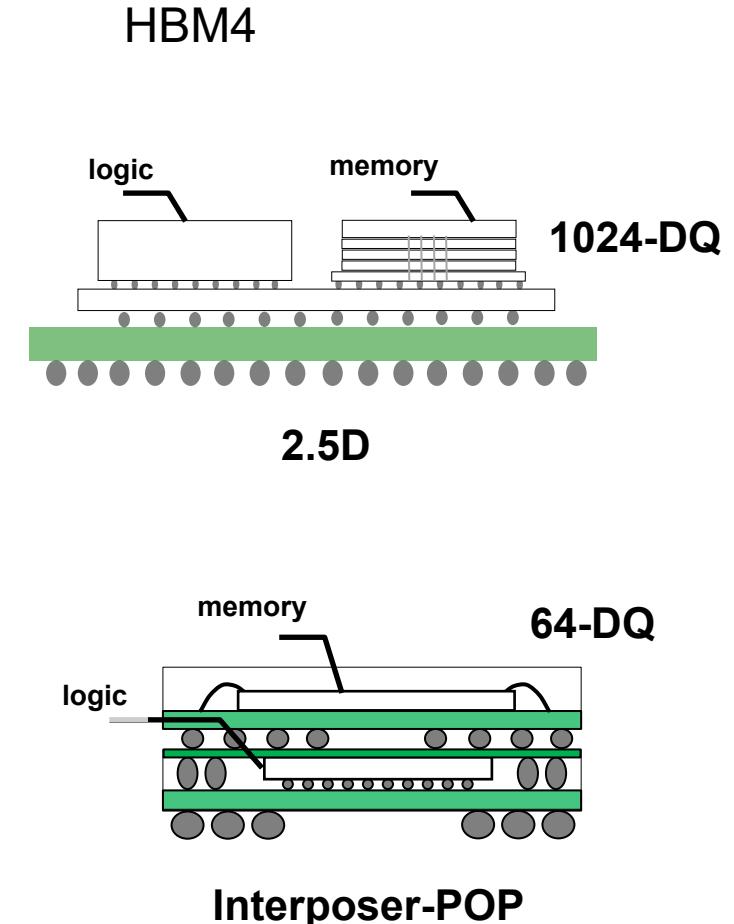
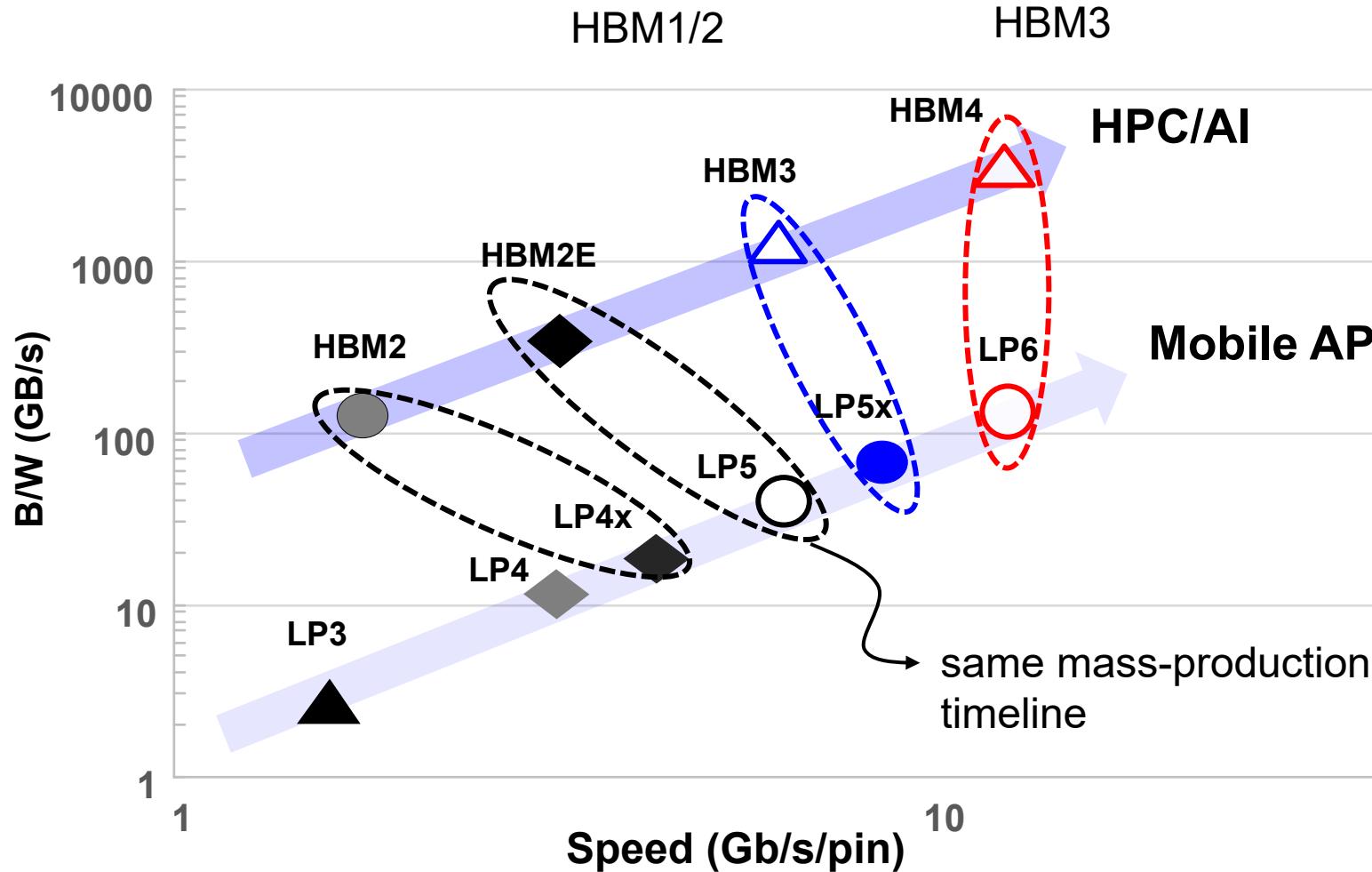
■ Implementation

■ Measurement Results

■ Conclusion

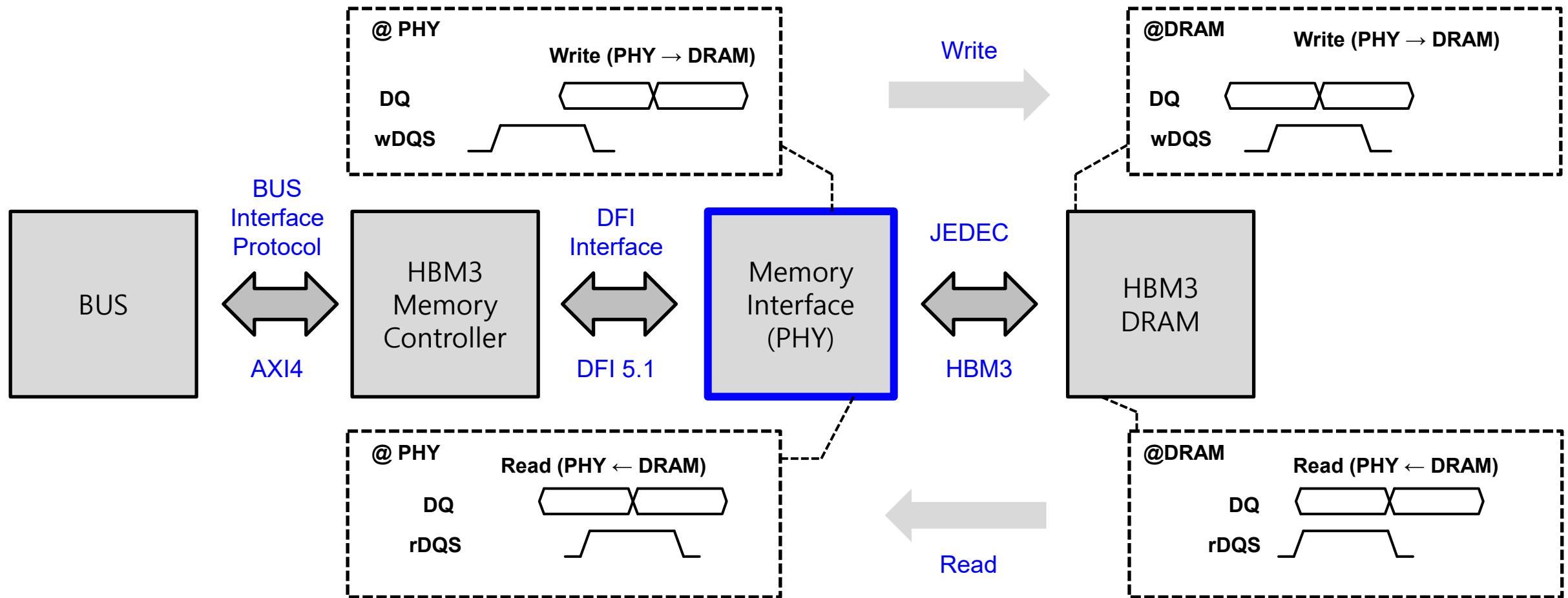
HBM Interface Trend

■ HBM Trend: Wide → Wide and Fast → Wider and Faster



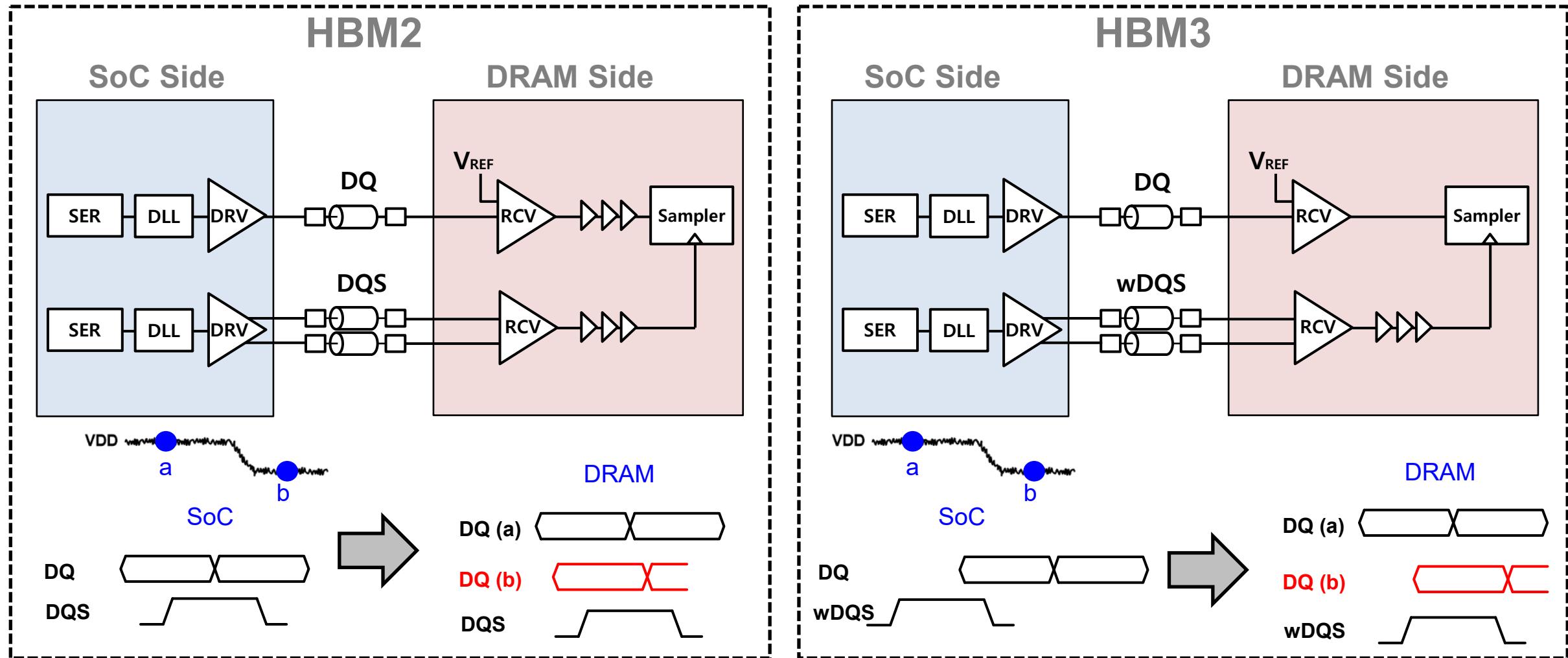
The Function of Memory Interface

- Memory I/F ensures reliable DRAM access



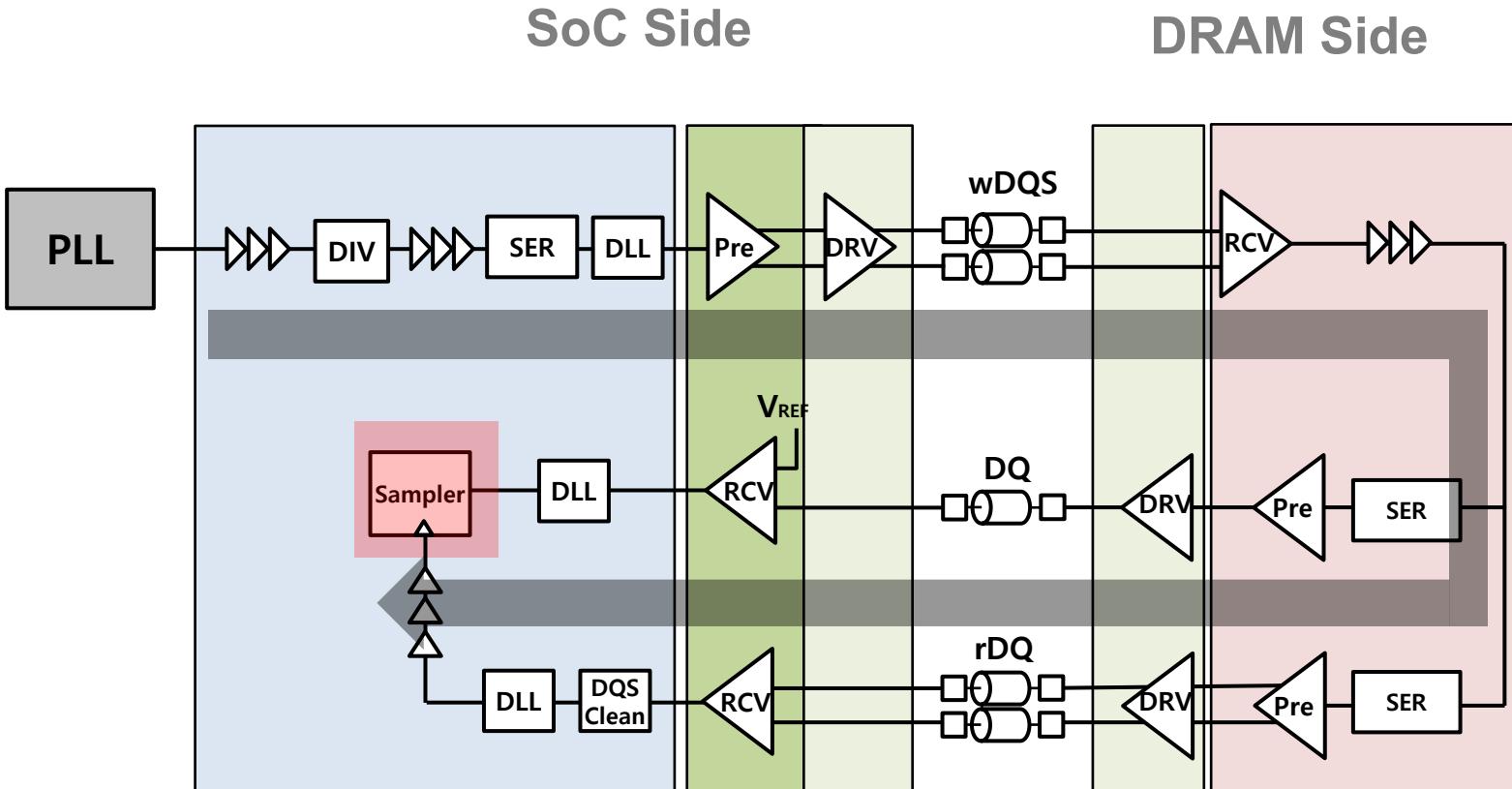
Supply Noise Sensitivity

- Unmatched DQ-wDQS → Supply noise tolerance ↓

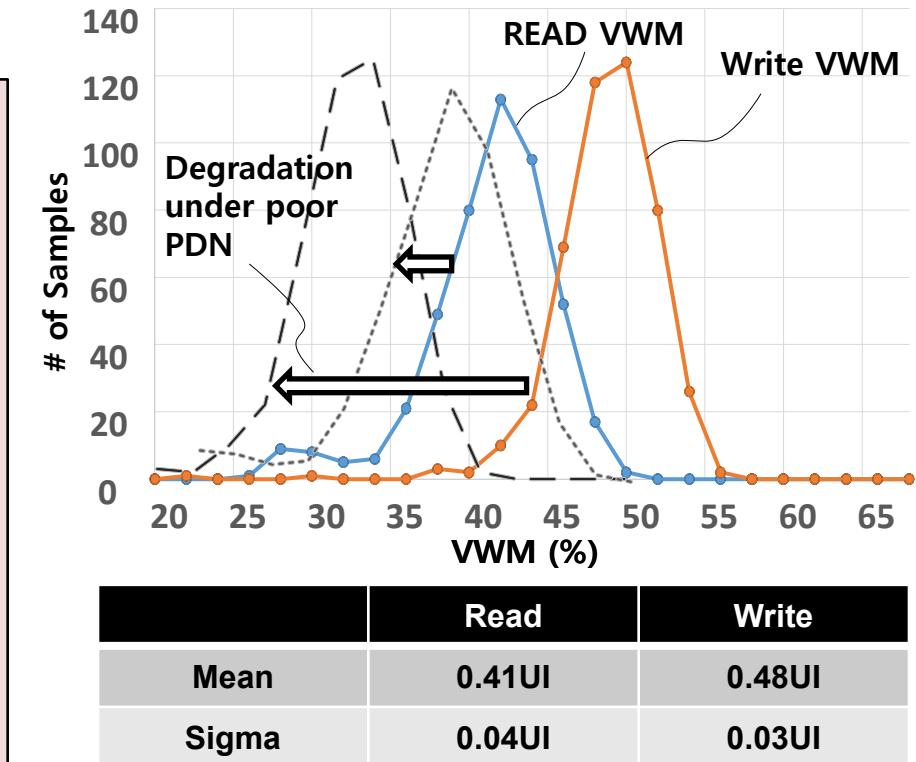


Read Path Characteristics

- Long turn-around path → Poor read VWM



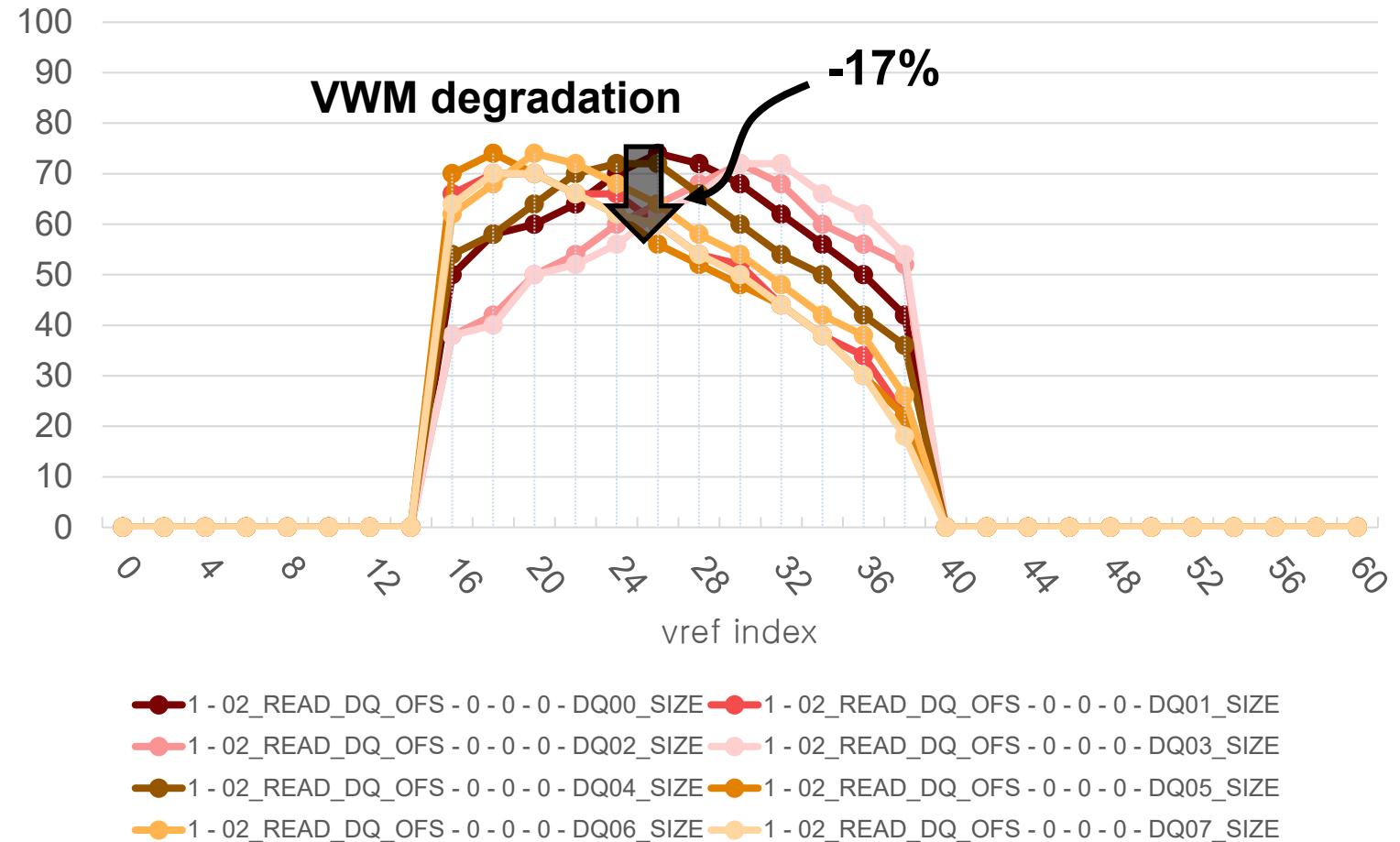
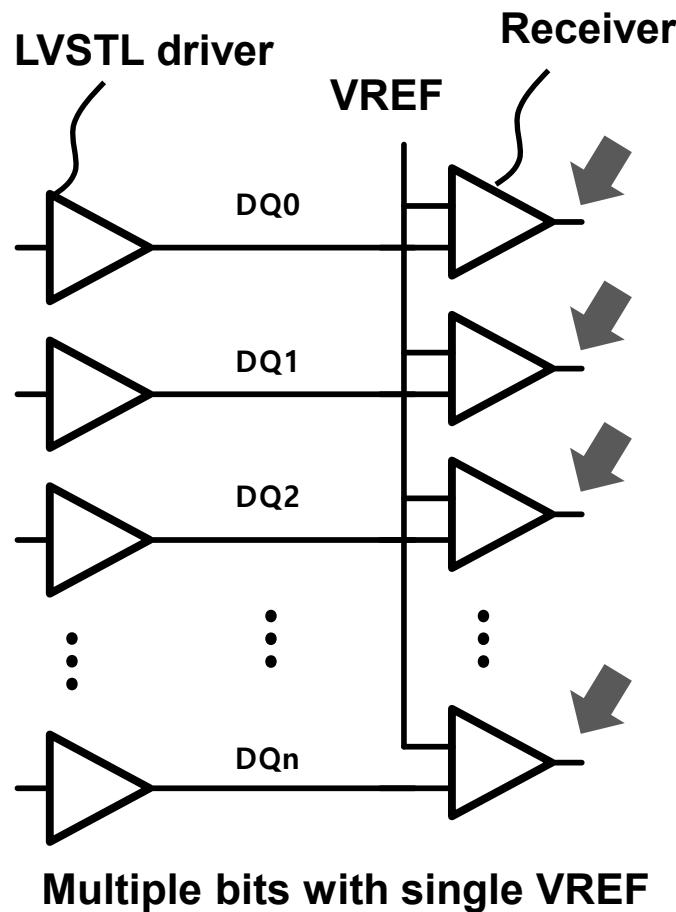
* Valid Window Margin



*Measurements results from LPDDR5x (WCK)

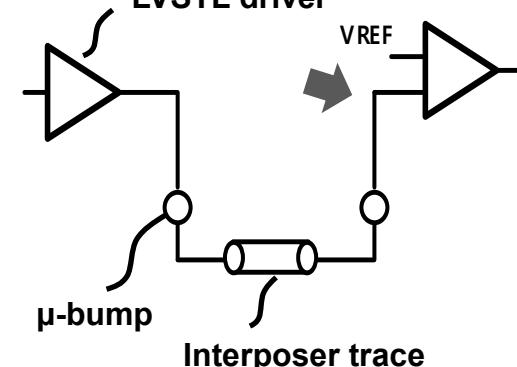
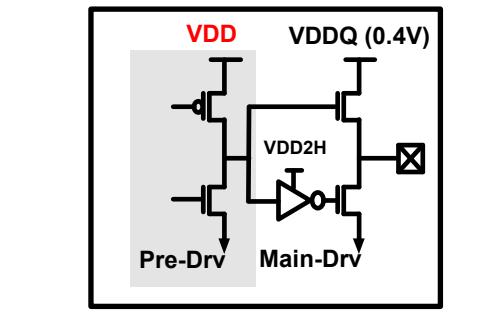
VWM Degradation

- Read VWM is significantly degraded by the input mismatch

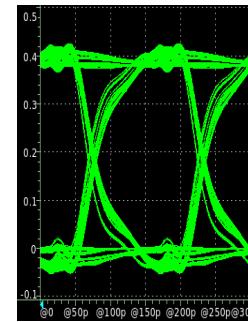


Channel Length

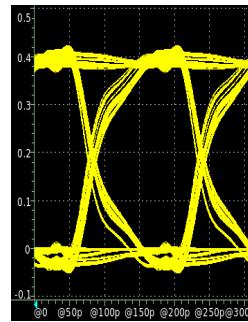
- The interposer channel length affects VWM performance
- PHY size impacts on the channel length



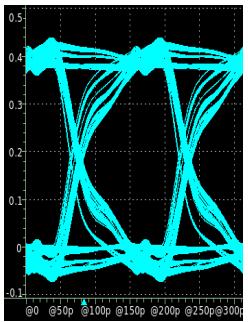
Length: 1000um ~ 6000um



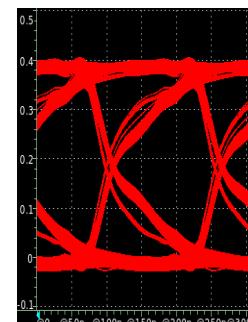
1000um



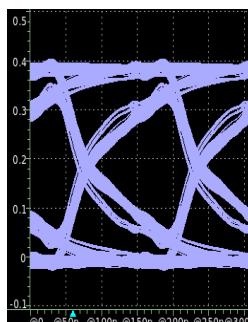
2000um



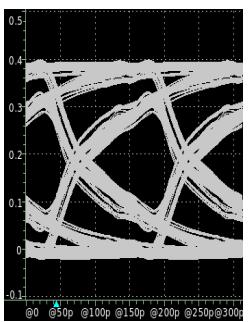
3000um



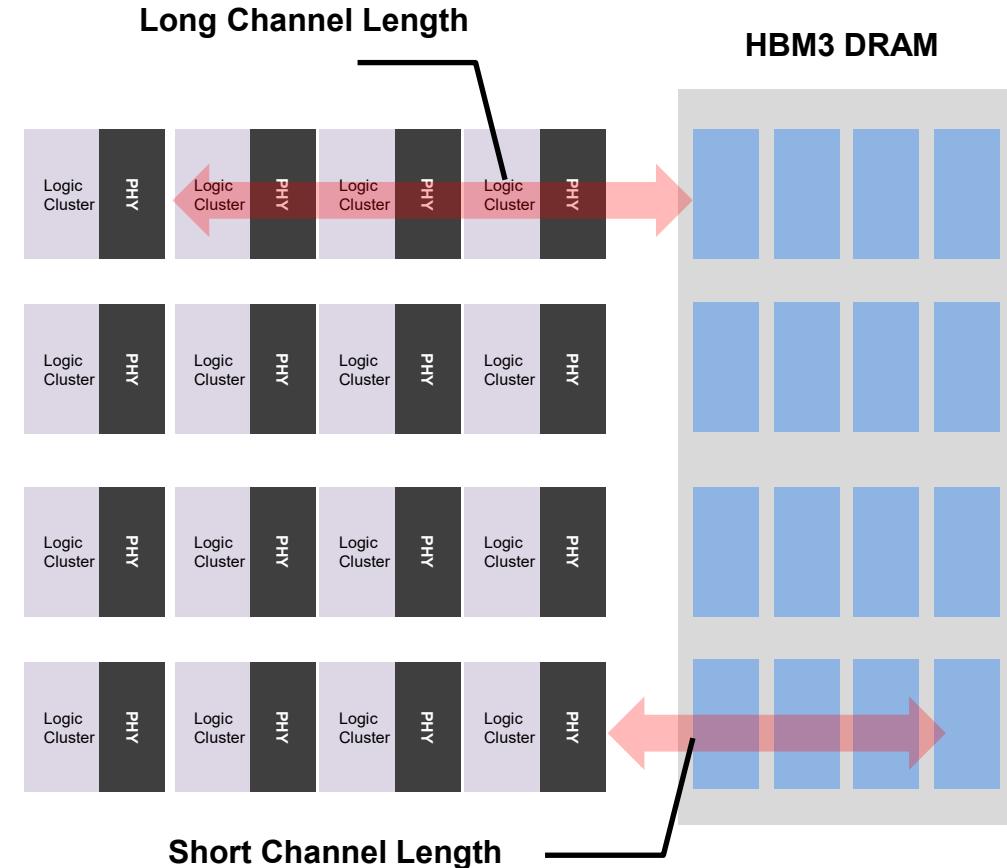
4000um



5000um



6000um



Outline

■ Introduction

■ Proposed Architecture

- Digital PHY Architecture
- TX and RX I/O Architecture
- Delay Compensation
- Real-Time Read Margin Detection

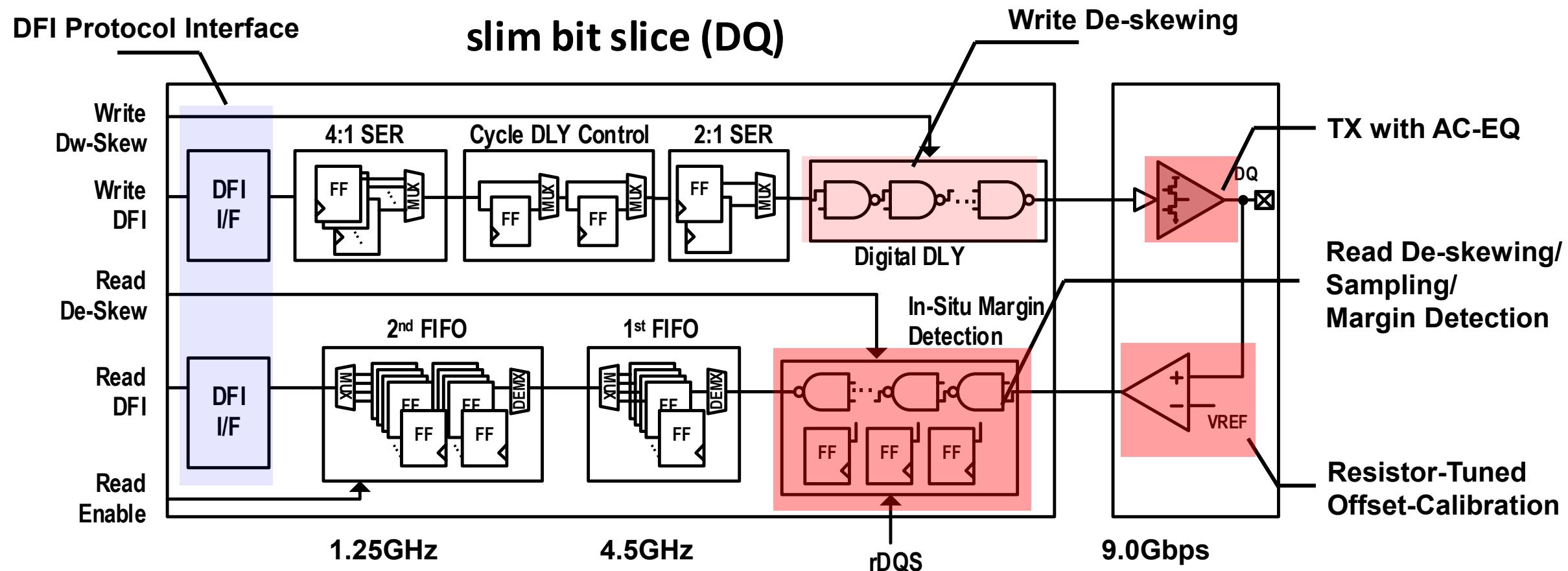
■ Implementation

■ Measurement Results

■ Conclusion

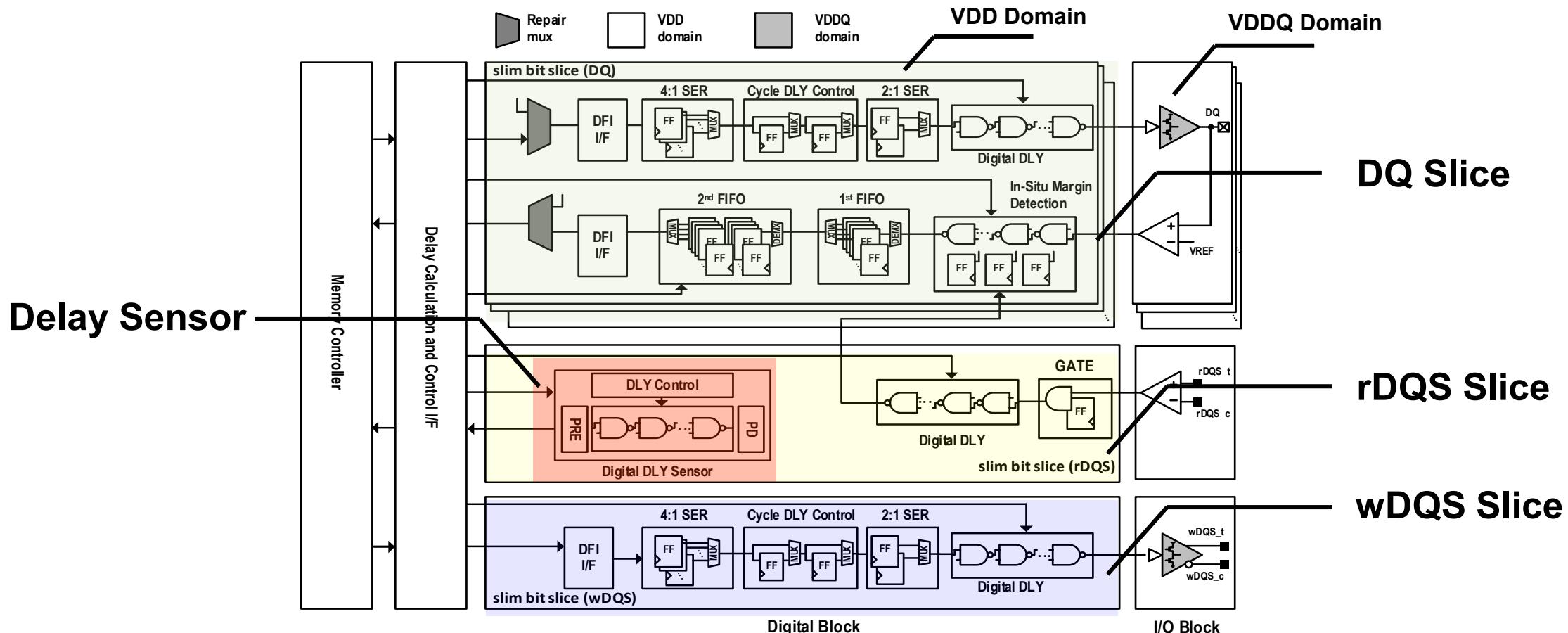
Digital Bit-Slice

- Slice-based digital PHY architecture
- Standard cells with logic partitioning to maximize PPA



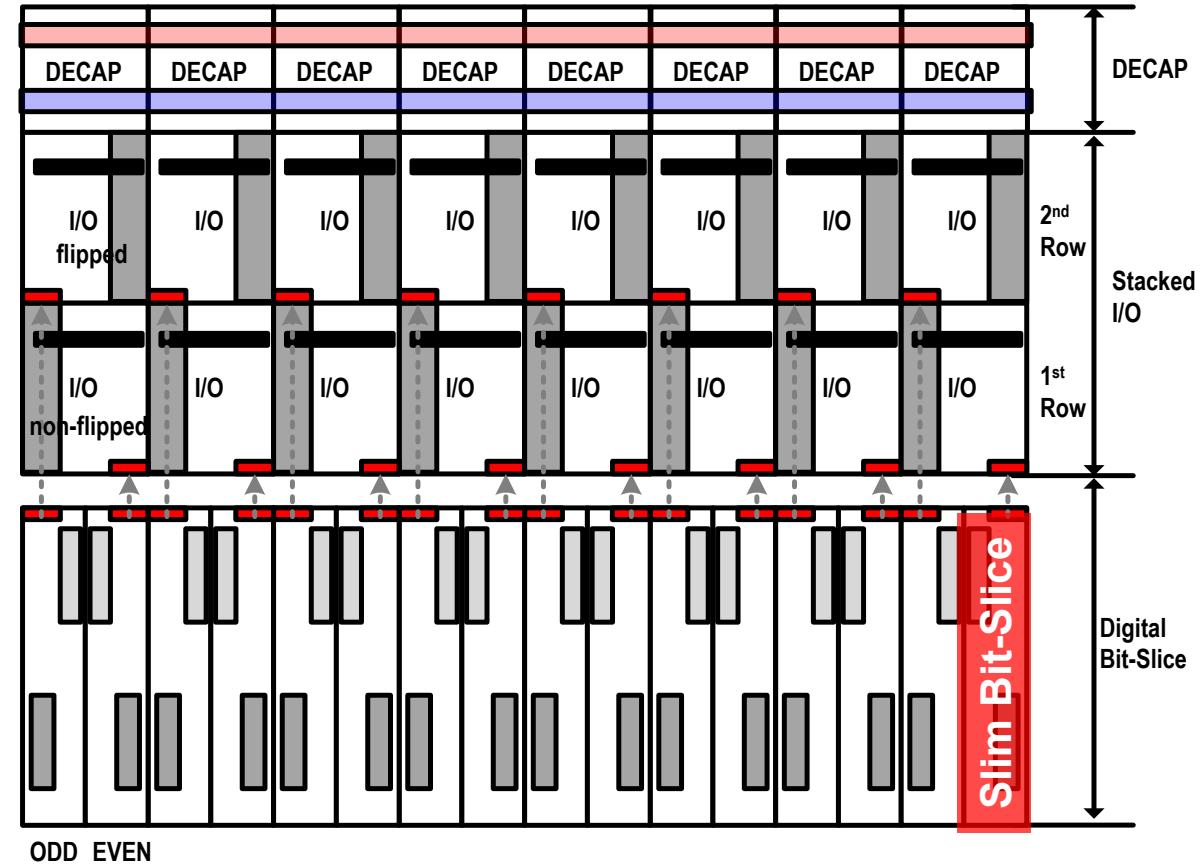
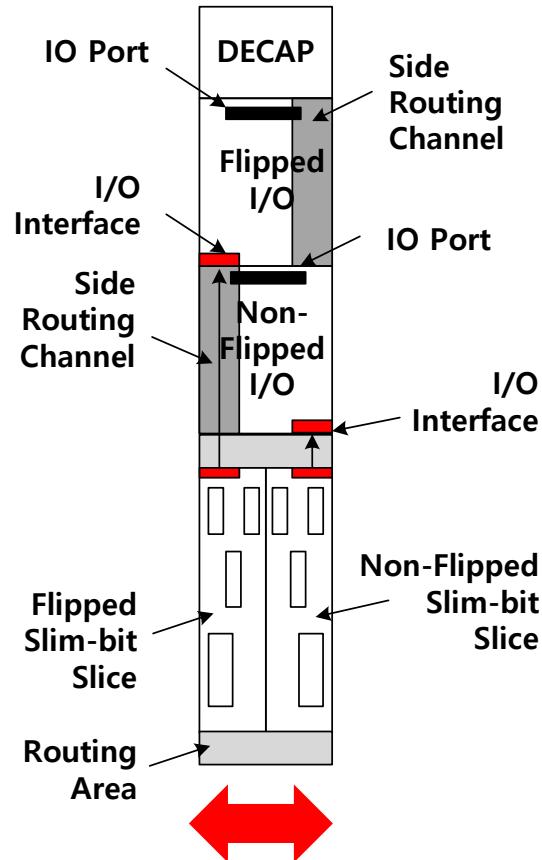
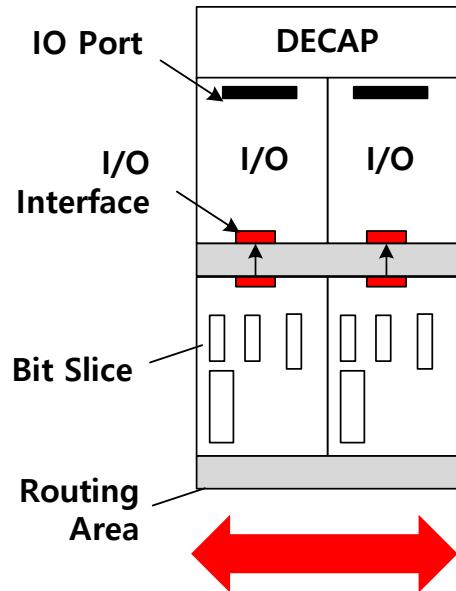
Digital PHY Architecture

- 3 types are used to constitute data PHY
- 2 voltage domains: VDD for digital and VDDQ for I/O



Slim Bit-Slice for I/O Stacking

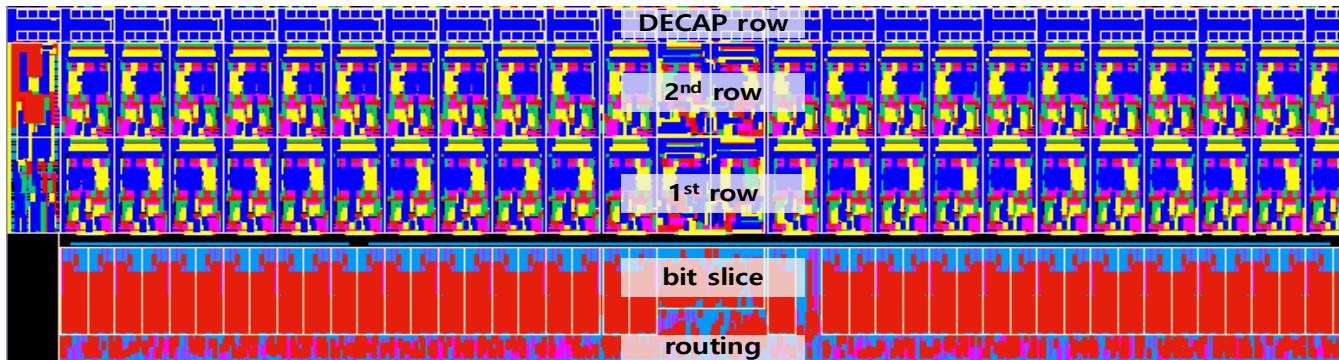
- Side-routing channel in I/O enables I/O stacking
- I/O stacking and slim-bit reduces the width by half



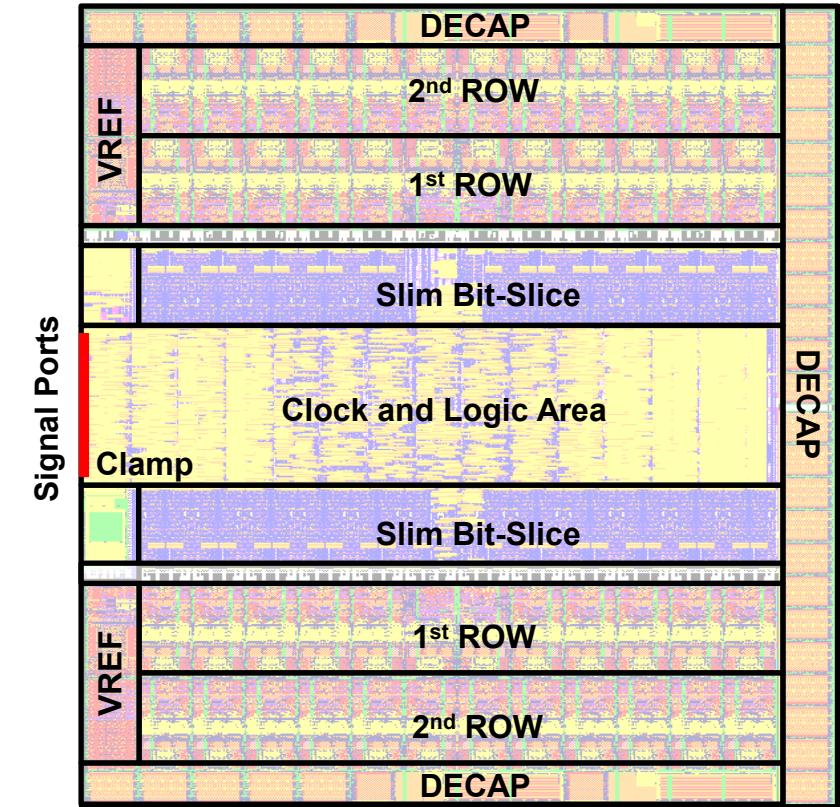
Folded PHY Structure

- Folded PHY combined with stacked I/O
→ 77.5% width reduction

Conventional Structure

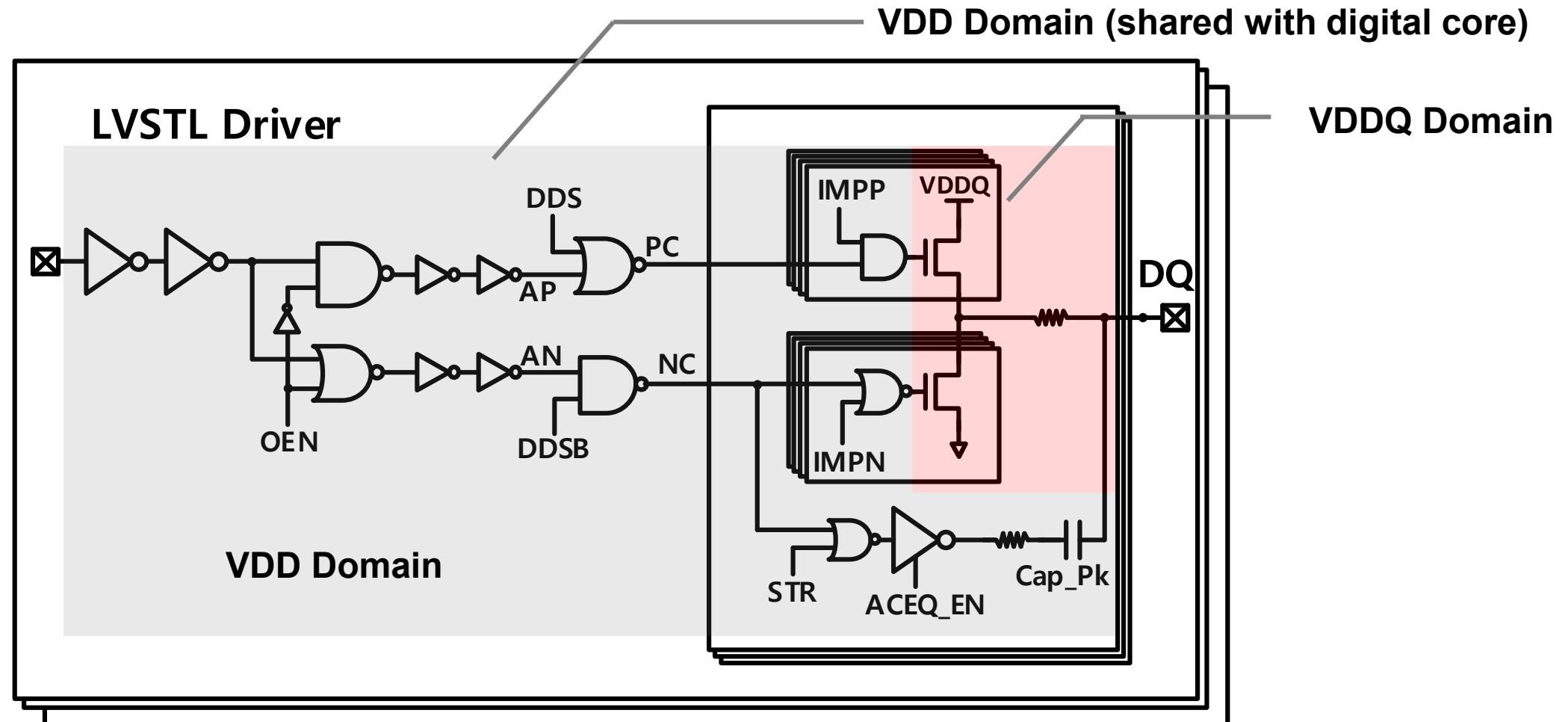


Folded PHY Structure (32b)



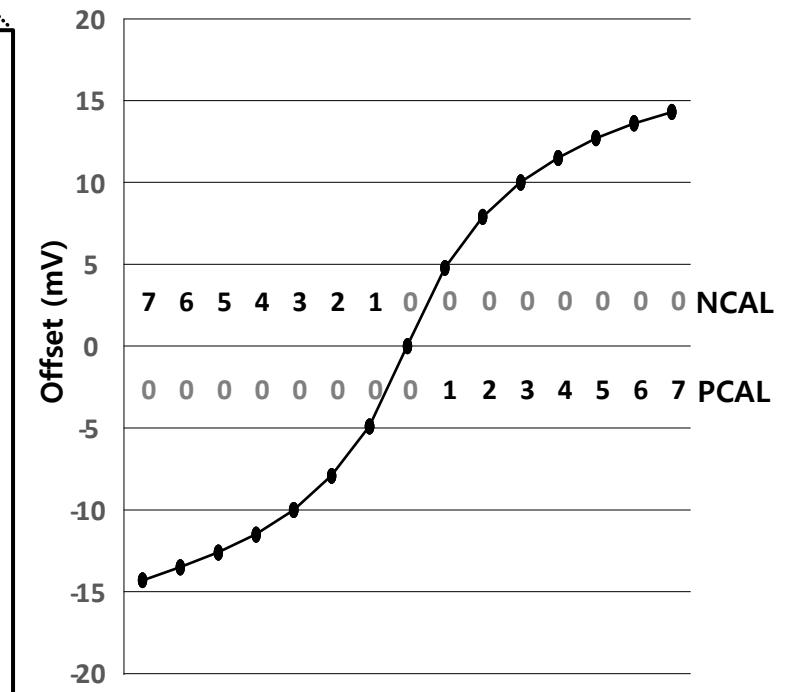
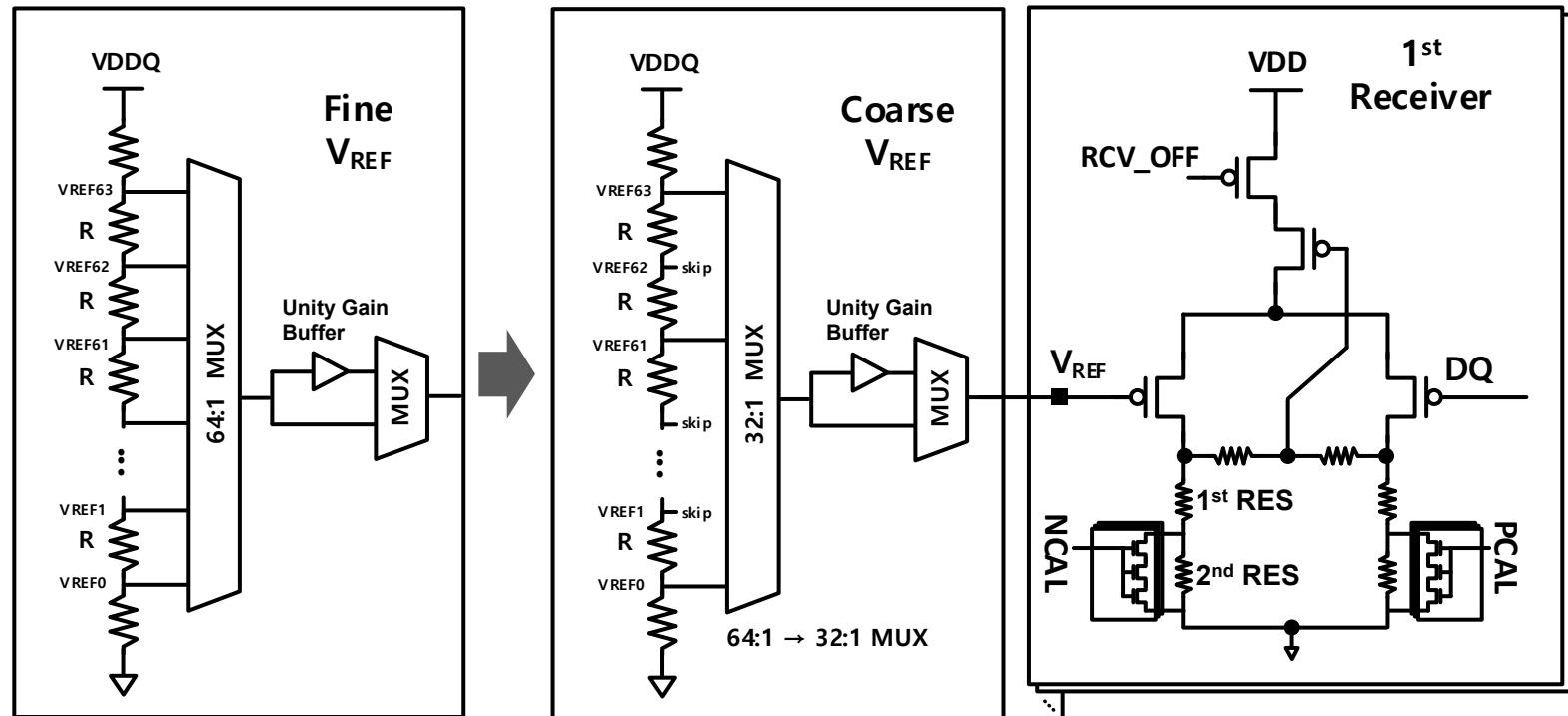
TX I/O with AC-EQ

■ LVSTL driver with AC-EQ to improve write VWM



RX I/O with Resistor-Tuned Offset-Calibration

- Fine V_{REF} + Offset-Cal. → Coarse V_{REF} + Offset-Cal.
 - Improved area efficiency
- Resistor-tuned offset calibration
 - Minimized B/W degradation

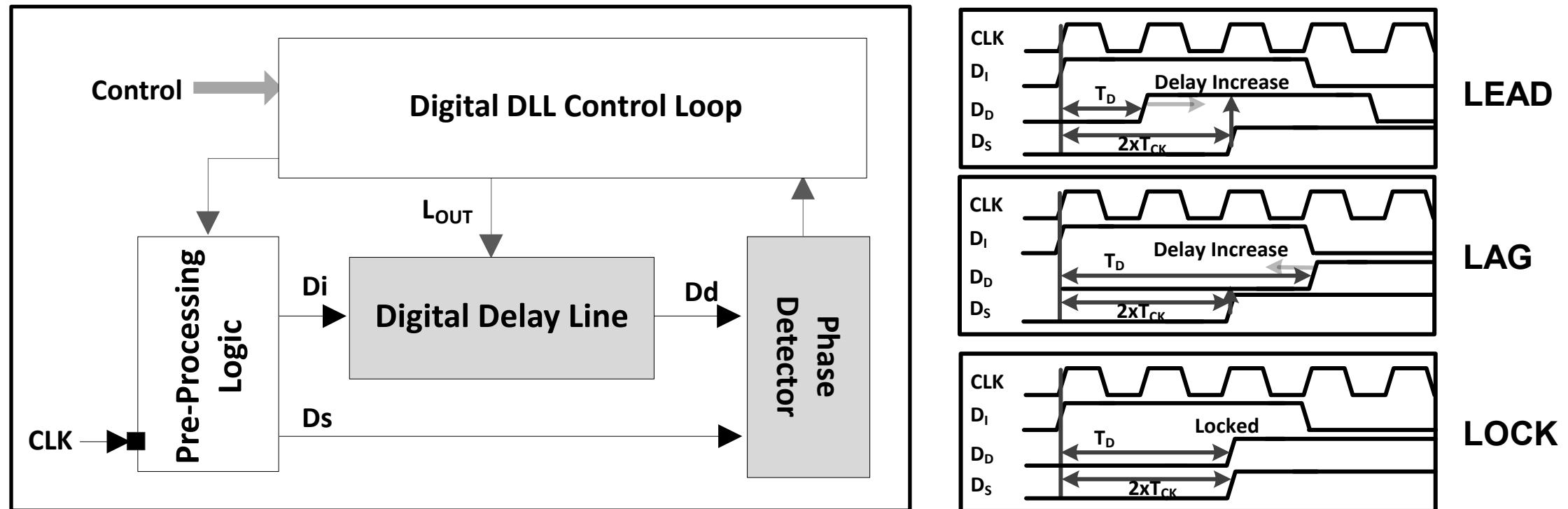


DLL-based Digital Delay Sensor

■ Digital DLL with the pre-processing logic (2-tCK Lock)

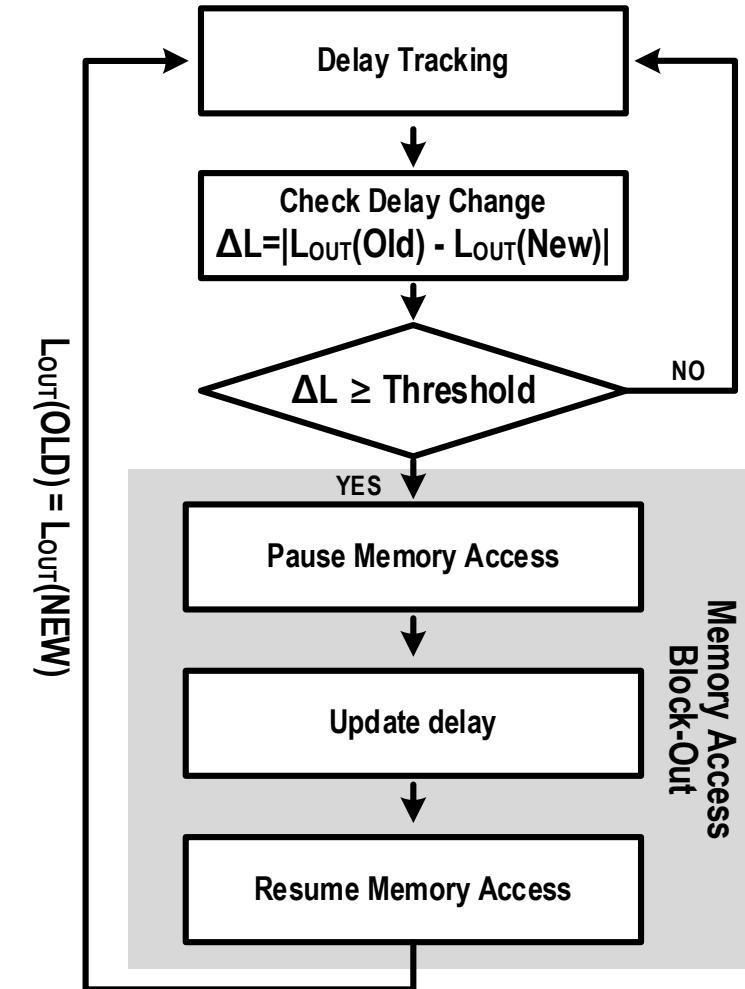
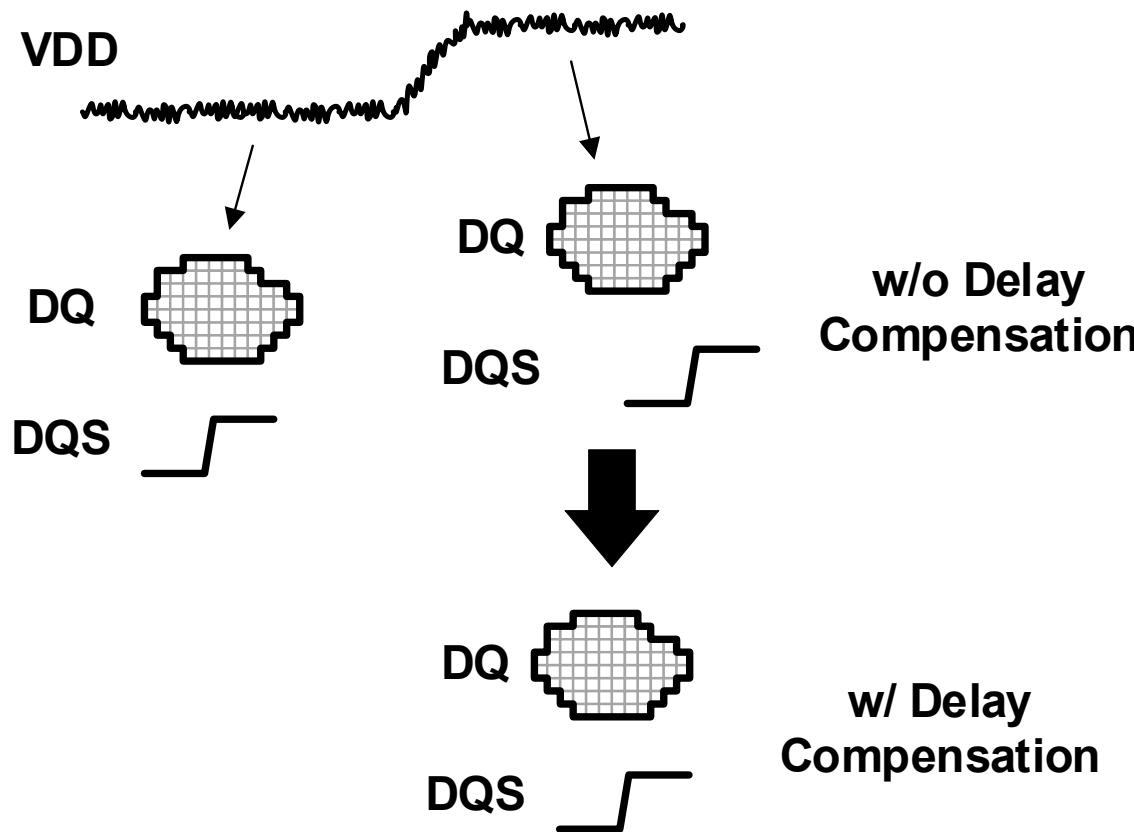
- Key Benefits: 1) simple control logic, 2) minimized harmonic-lock risk, 3) minimizes dynamic power, 4) increased sensitivity, 5) minimized error

Total error = (phase error + linearity error)/(# of clock cycles)



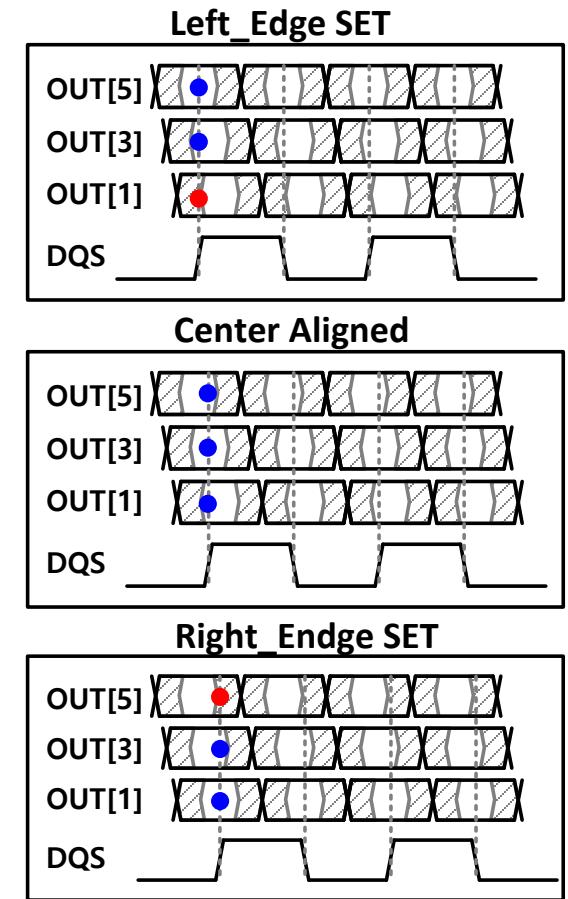
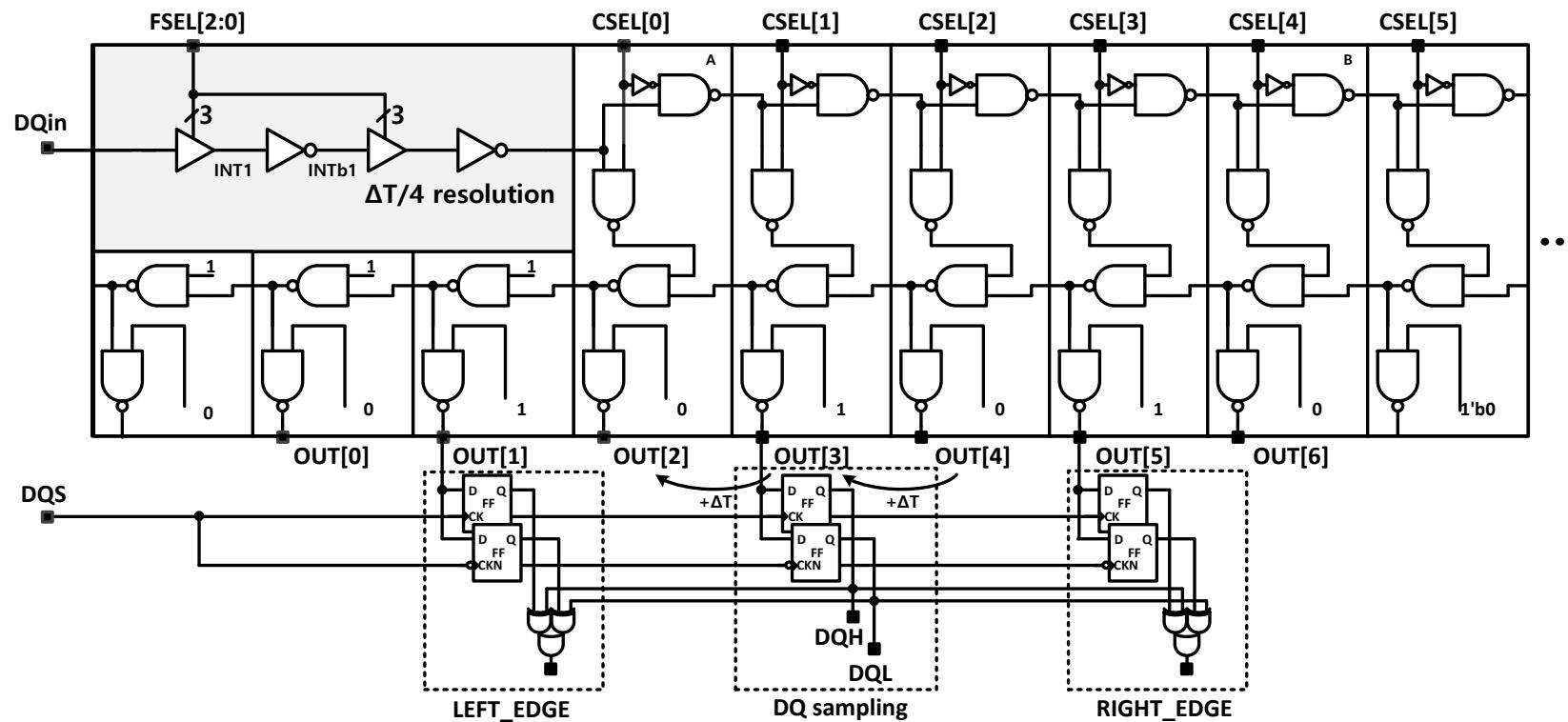
Delay Compensation with Delay Sensor

- Delay sensor detects voltage change to compensate for delay variation



Read Margin Detection

- Delay line for de-skewing includes discrepancy detection
- Read valid-window-margin is detected from real-time data
- Periodic training is not required



Outline

■ Introduction

■ Proposed Architecture

■ Implementation

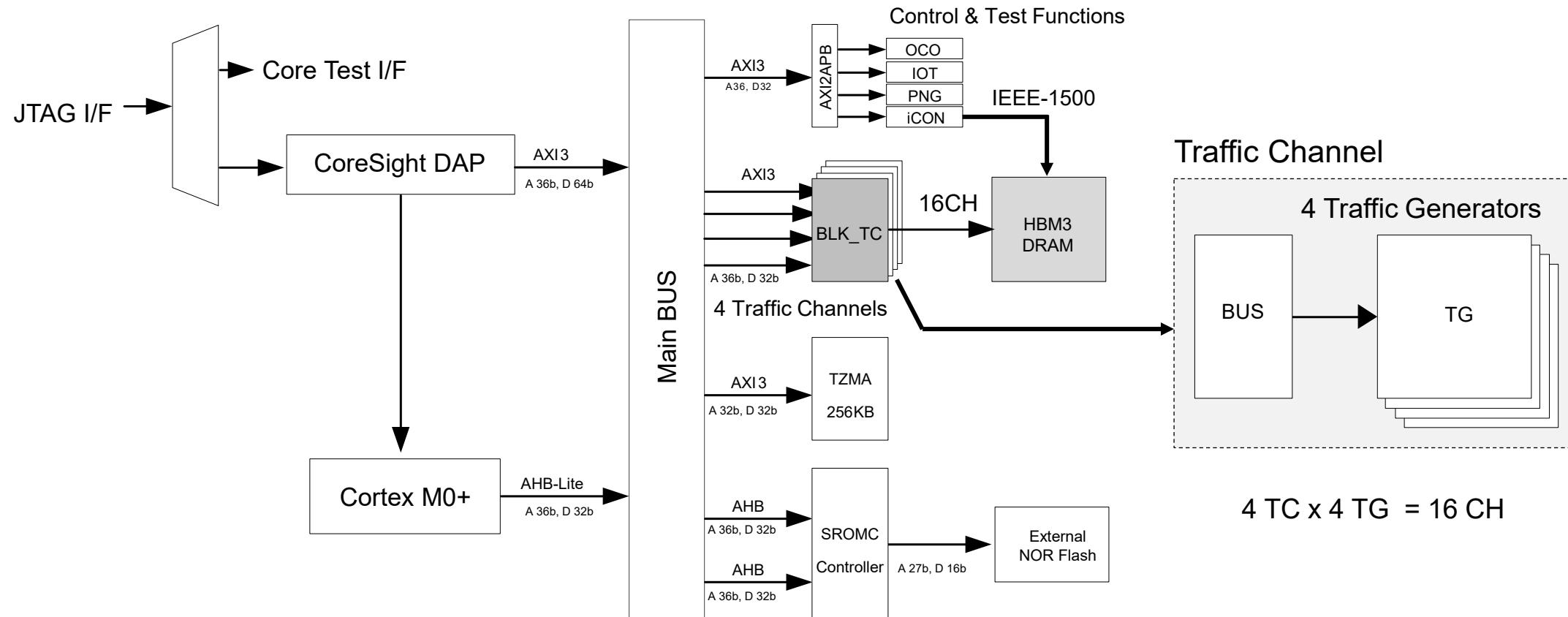
- Test-Chip Architecture
- PHY and Test-Chip Implementation

■ Measurement Results

■ Conclusion

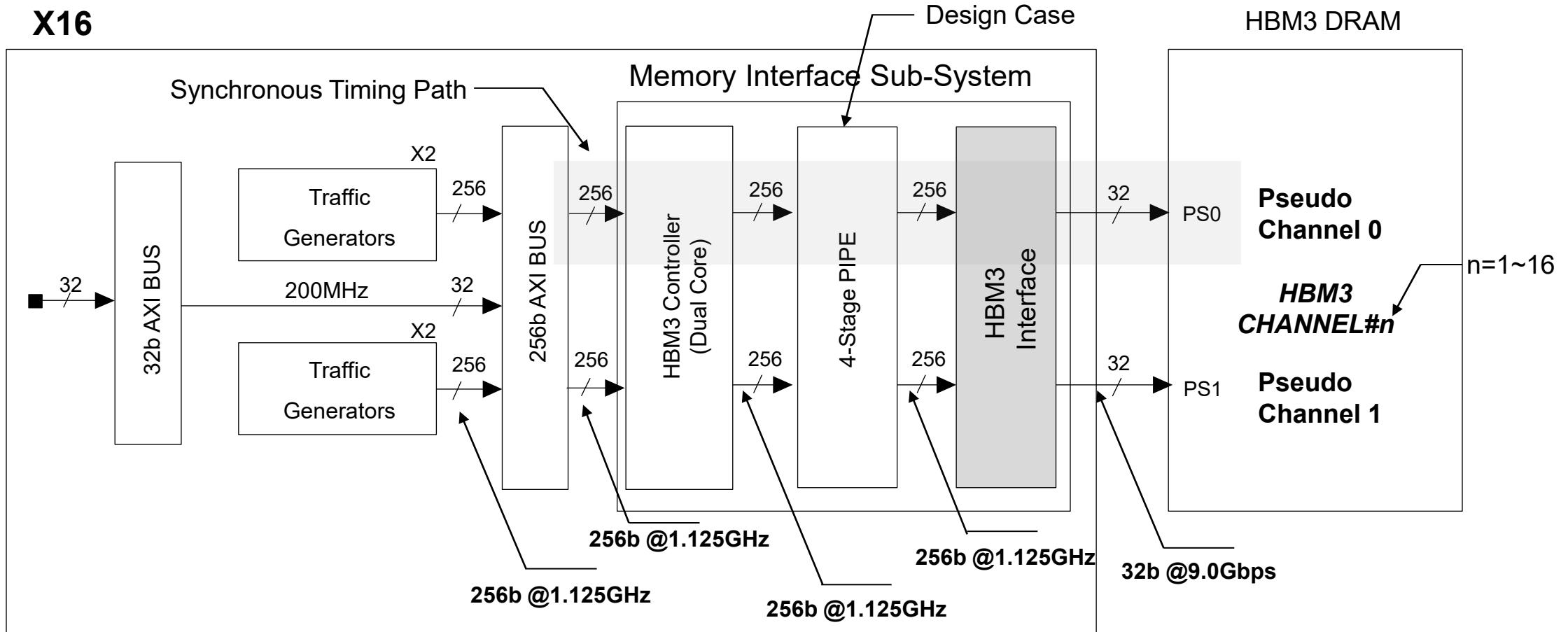
Test-Chip Architecture

- Architecture for heavy-traffic generation
- Local traffic generation considering implementation



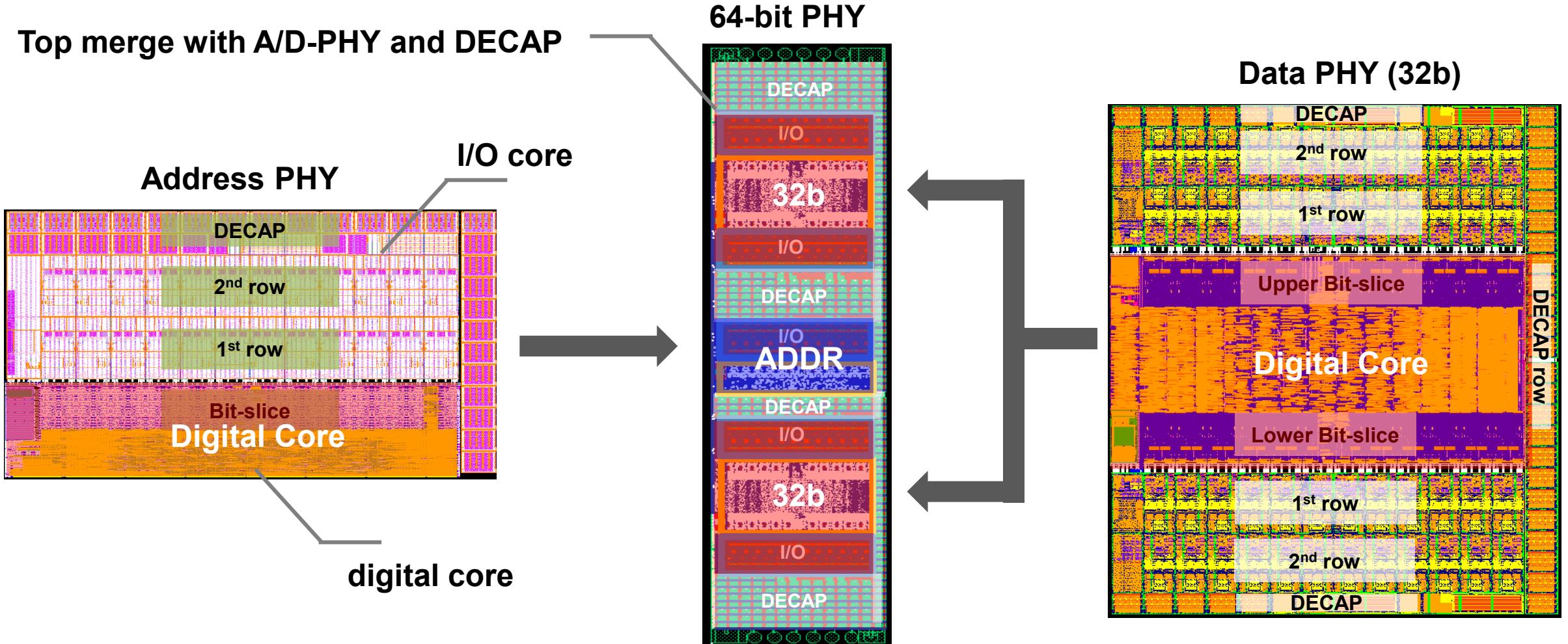
Traffic Generator Block

- Four 256b DMAs per channel for heavy traffic generation
- 1.125GHz sync. timing in memory interface sub-system



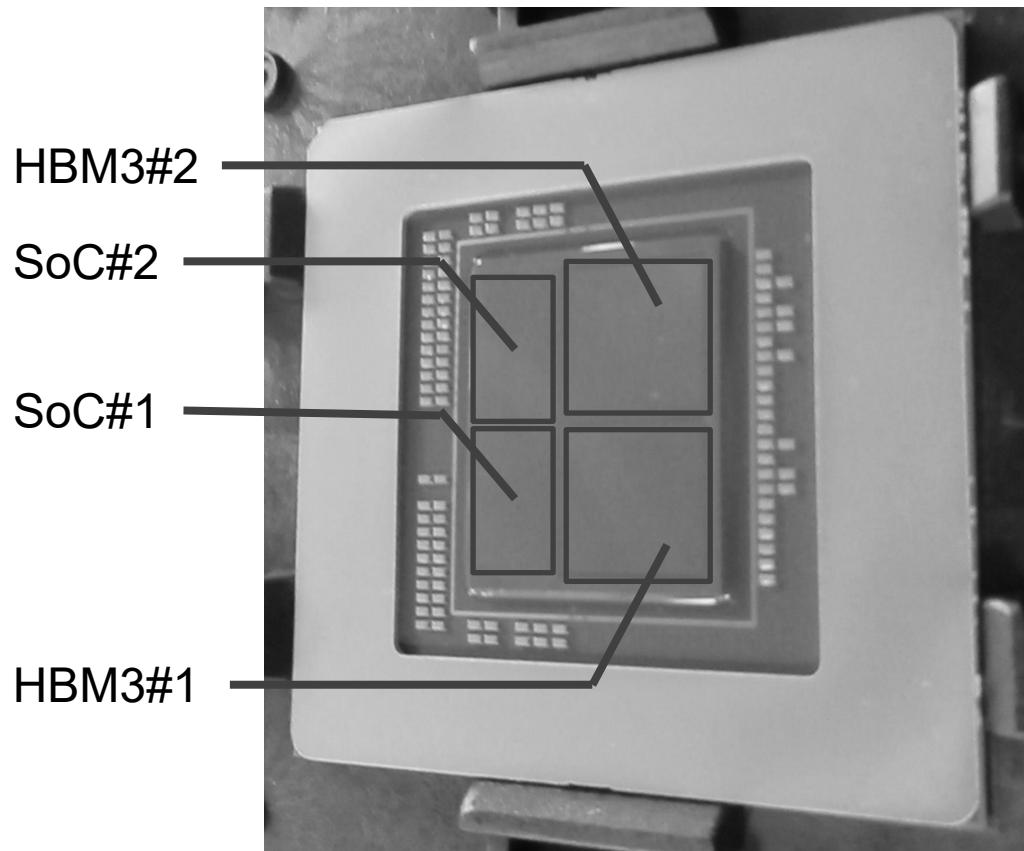
Auto P&R-Based PHY Design

■ Hierarchical design based on digital design methodology



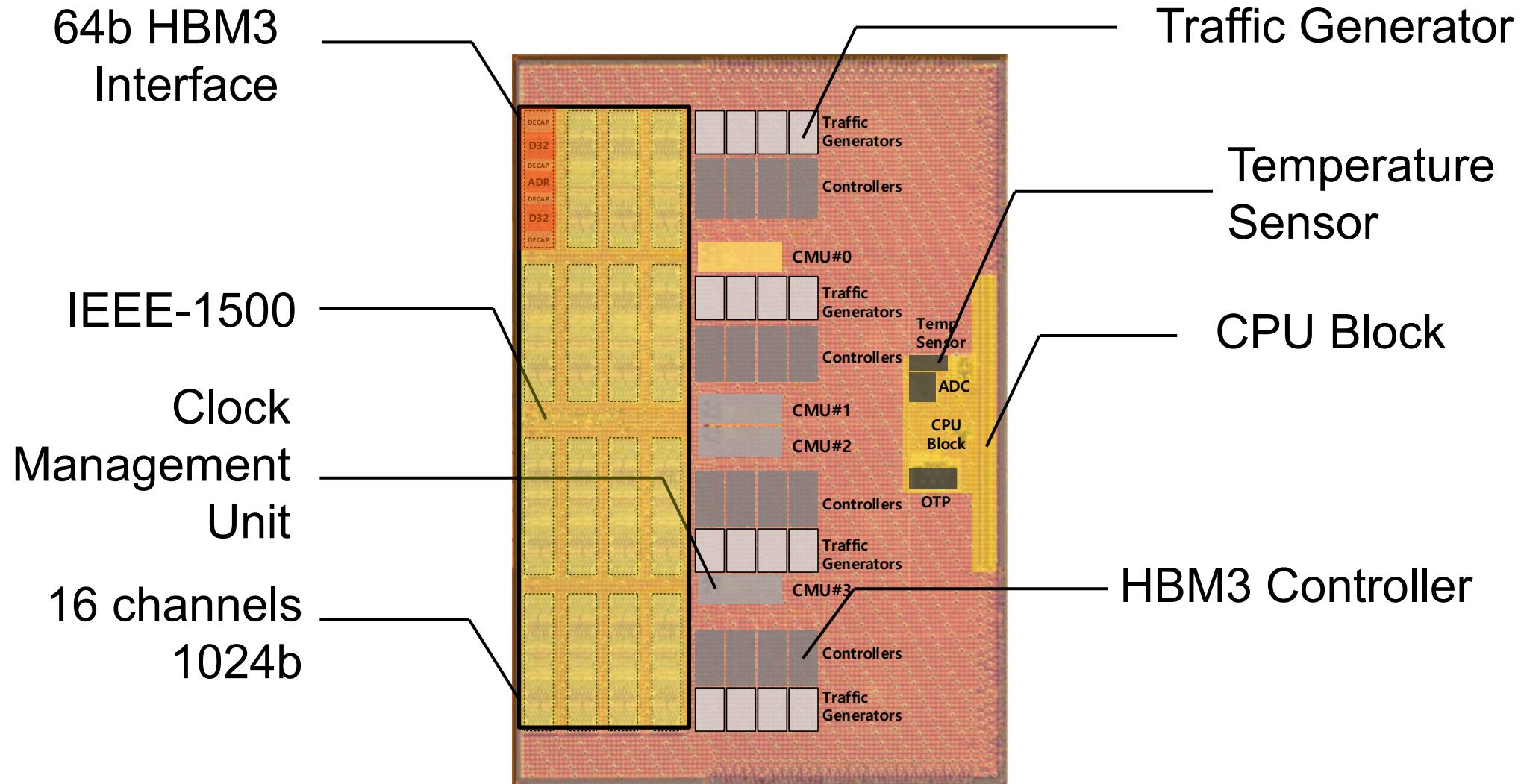
Test-System Summary

- The package includes 2 test-chips, 2 HBM3-cubes and interposer

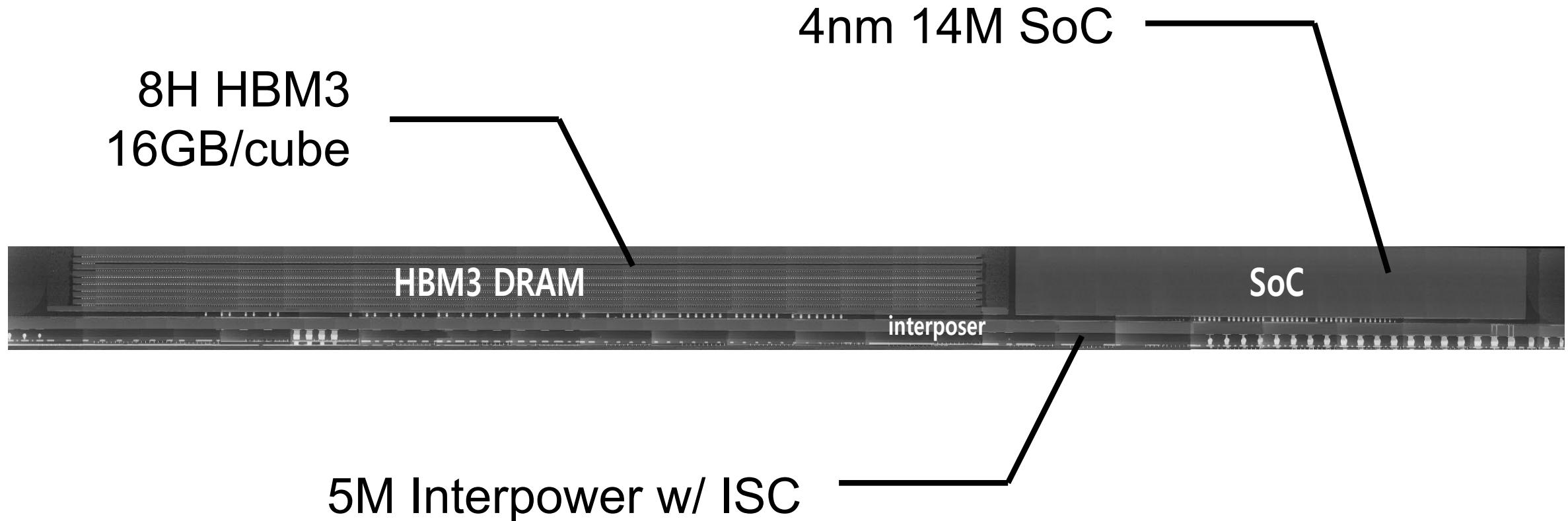


Process	4nm FinFET
Metal Stack	14M
DRAM	8H HBM3
DRAM Die Density	16Gb/die
Capacity	16GB/cube x 2 (32GB)
Max. B/W per cube	1.15TB/s/cube
Interposer	5M Si. Interposer
Capacitor	ISC

Test-Chip Floorplan



2.5D-Package Cross Section

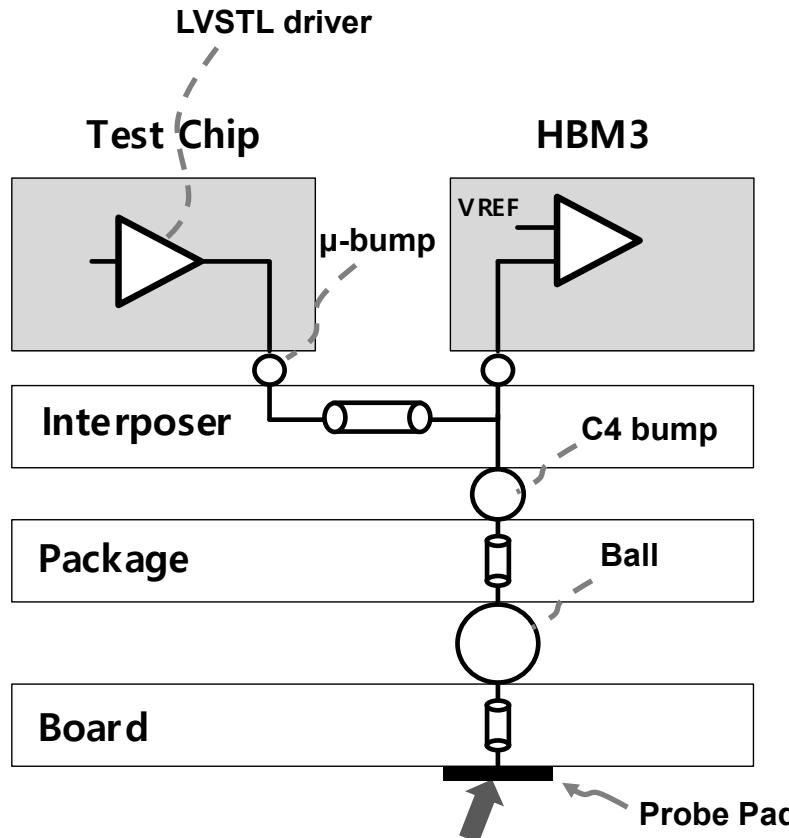


Outline

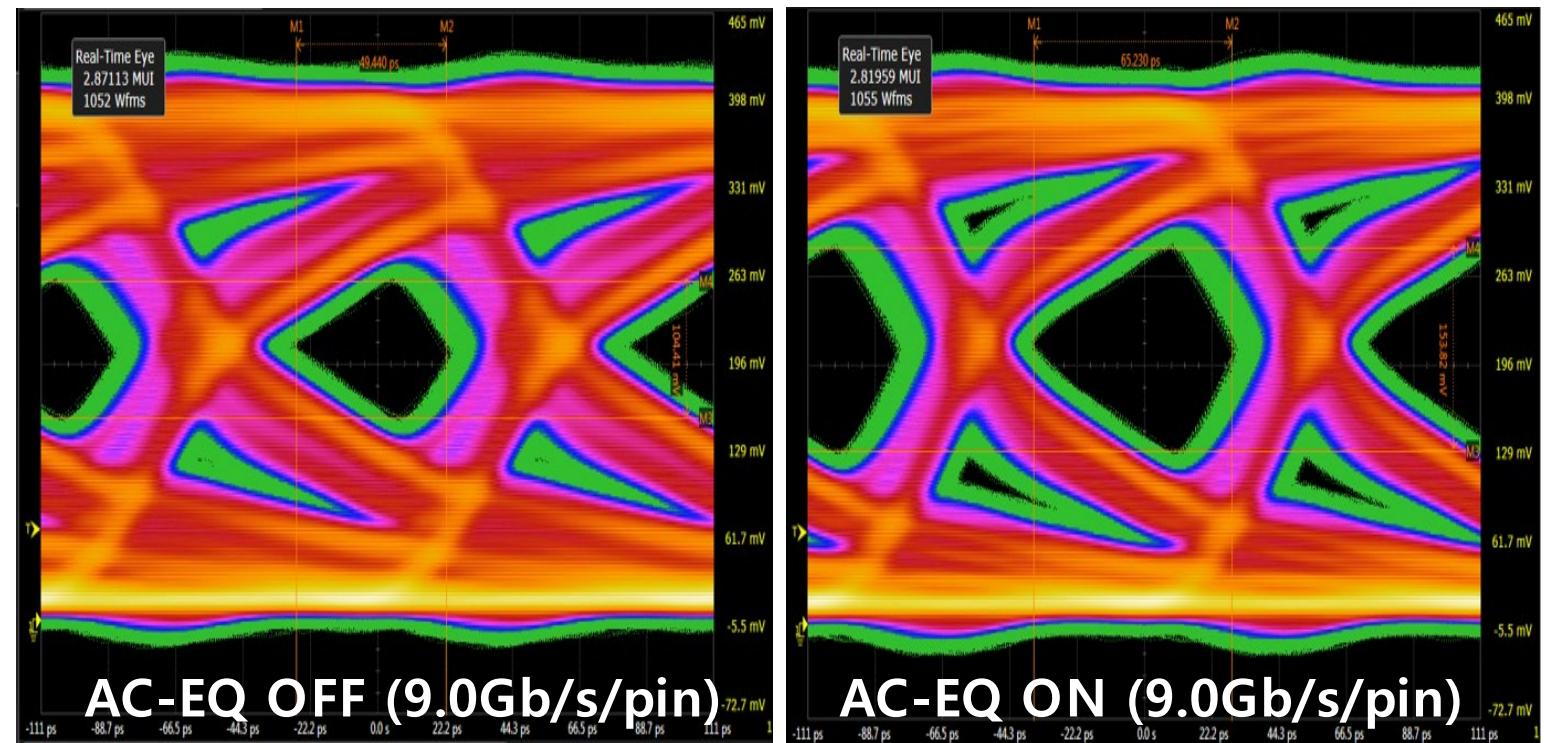
- Introduction
- Proposed Architecture
- Implementation
- **Measurement Results**
- Conclusion

Measured Waveform

- 9.0Gb/s/pin operation with 660mV VDD and 300mV VDDQ
- AC-EQ improves eye height up to 47%

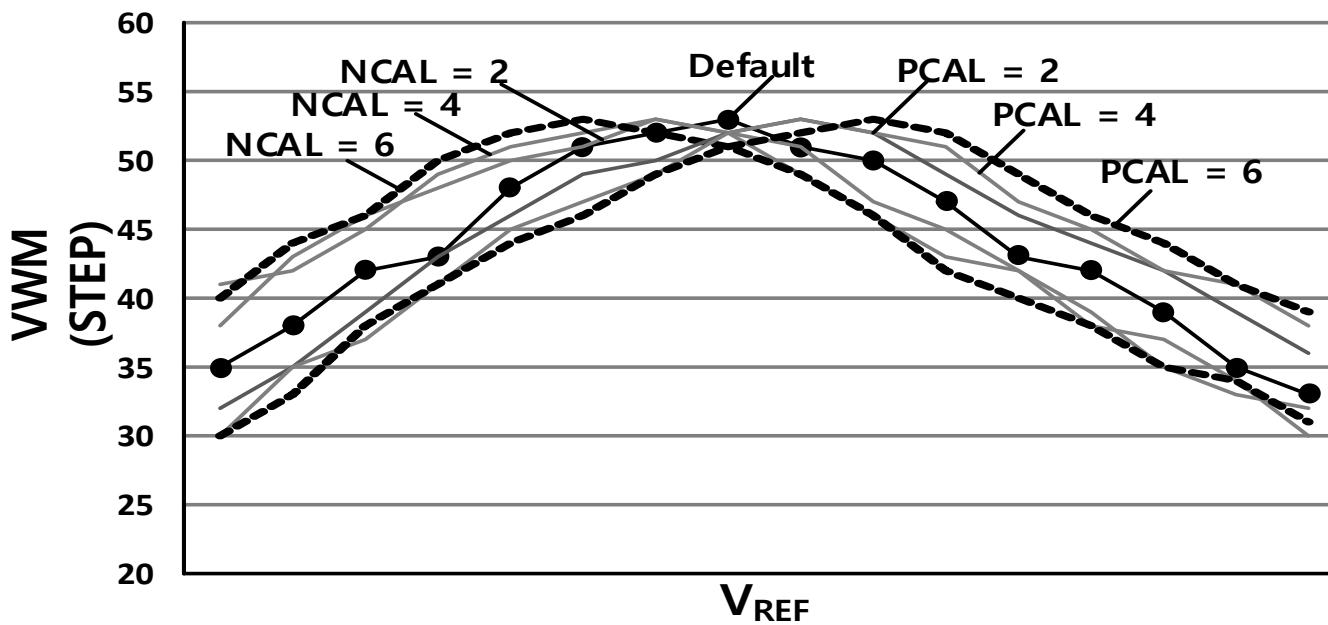


Measured at this point



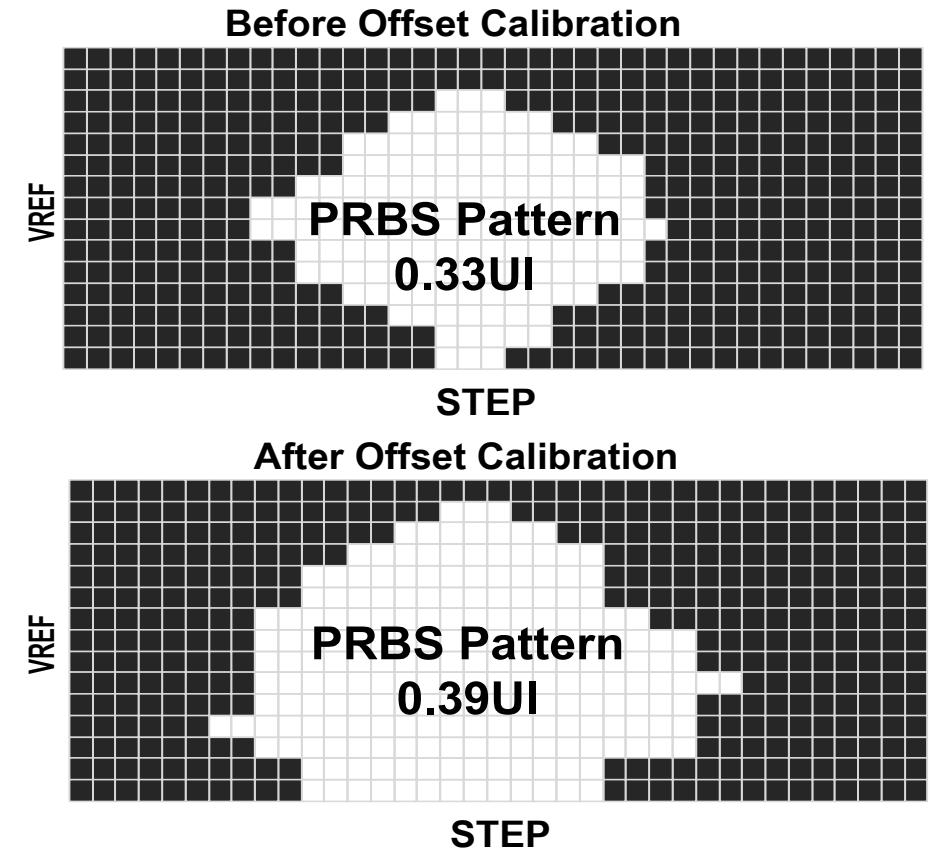
Measured Offset-Calibration

- Offset-Cal. effectively compensates for mismatches
- Read VWM is improved by 16.7%



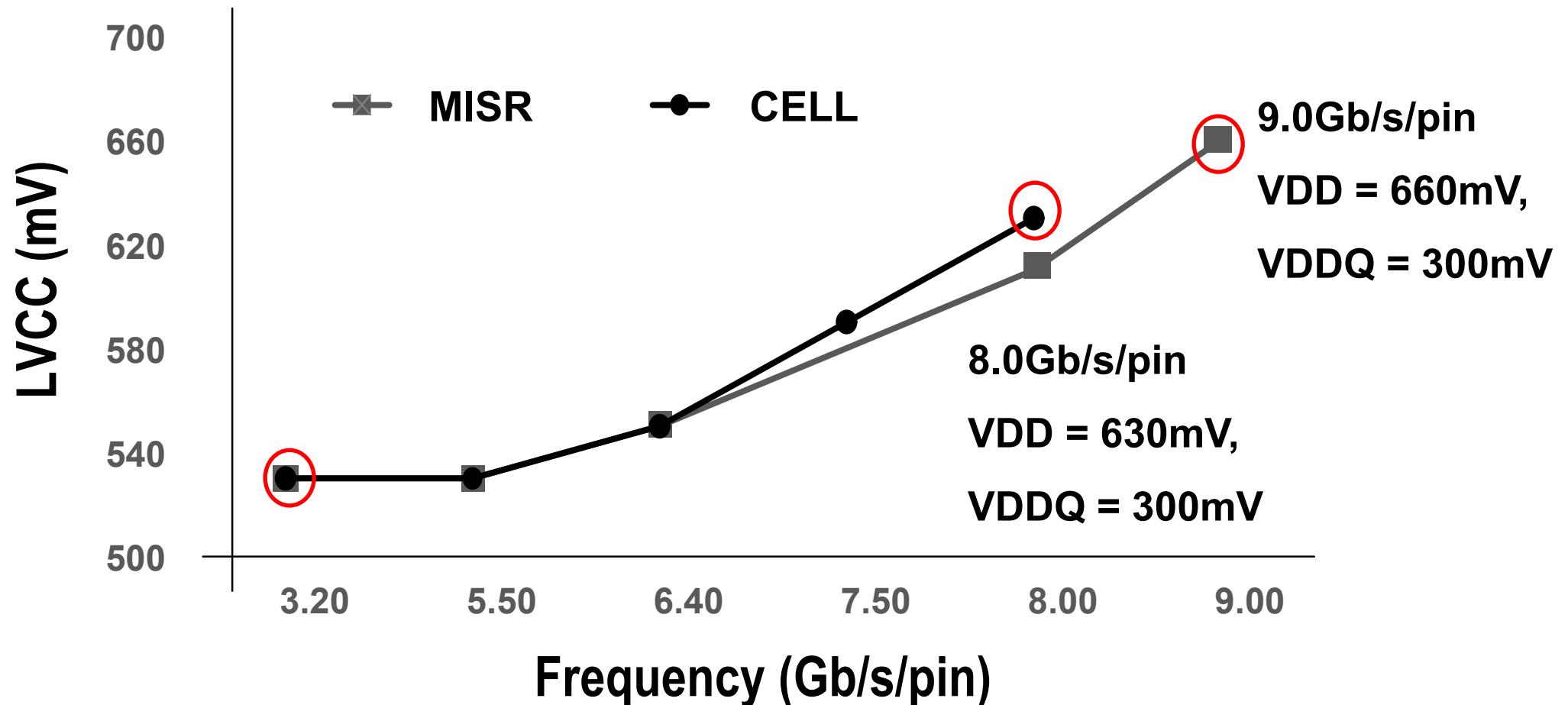
Coarse VREF + Offset Calibration
(Measurement)

64b channel eye diagram



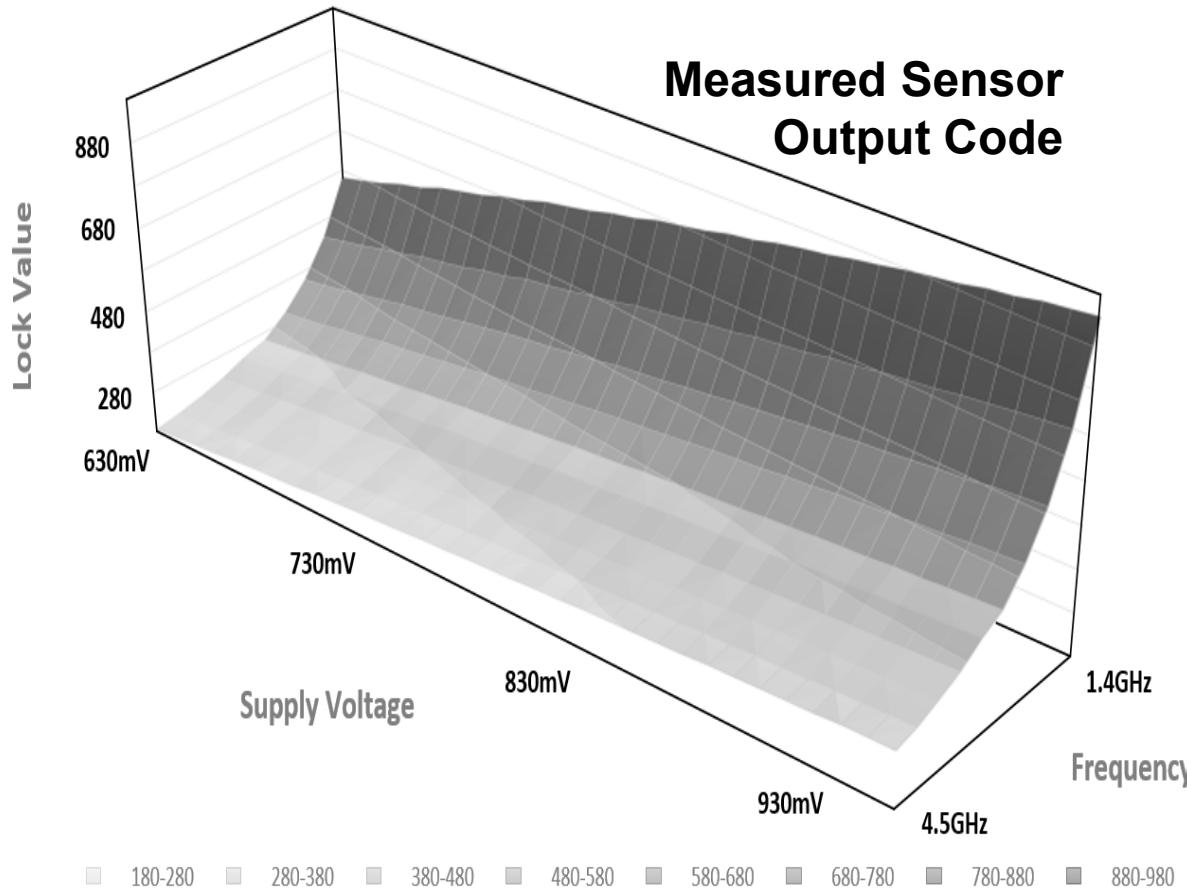
Measured Operating Range

- 8.0Gb/s/pin at 630mV (Cell), 9.0Gb/s/pin at 660mV (MISR)
- Achieved 3.2Gb/s/pin at 530mV

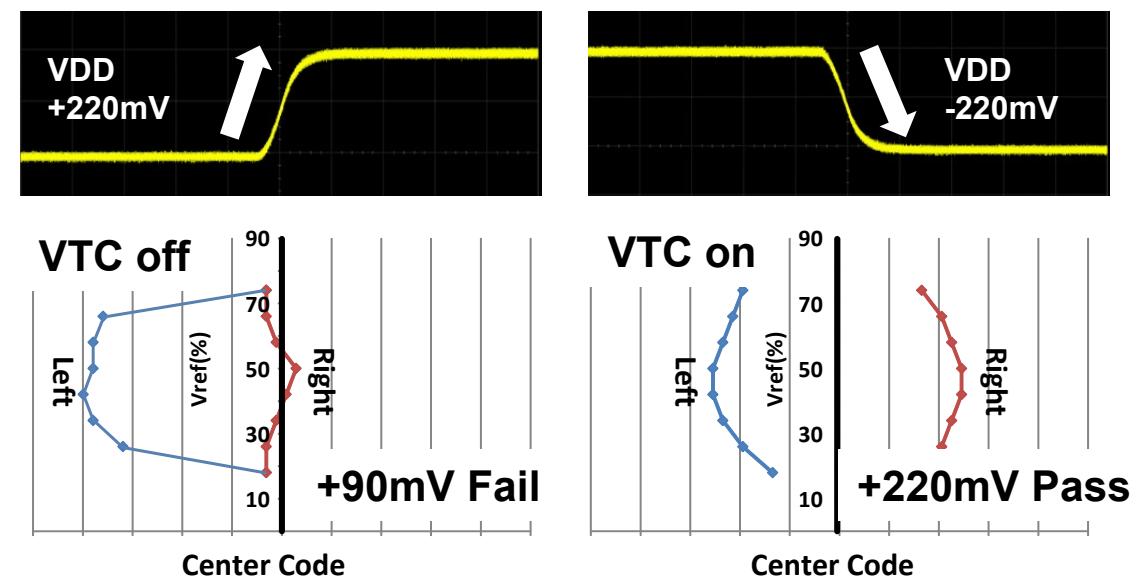


Measured Delay Sensor and Voltage Tolerance

- Wide operating range: 1.4GHz ~ 4.5GHz@620mV
- Supply noise tolerance: $\pm 220\text{mV}$



Tolerance	From	To
+ ΔV (+220mV)	700mV	920mV
- ΔV (-220mV)	920mV	700mV



Summary Table

Reference	VLSI`19	ISSCC`20	HC`19	VLSI`19	This Work
Interface	LPDDR5	GDDR6	HBM2E	HBM3 Receiver	HBM3
Technology	8nm	8nm	7nm	65nm	4nm
# of Signal (Data)	40 (16)	40 (16)	1696 (1024)	1 (1)	1920 (1024)
Speed	7.3Gb/s/pin	18Gb/s/pin	3.2Gb/s/pin	4.8Gb/s/pin	9Gb/s/pin (*)
B/W	14.6GB/s	36GB/s	409.6GB/s	N/A	1.15TB/s
VDD/VDDQ	0.79V/0.5V	0.85V/1.35V	0.75V/1.2V	1.1V	0.66V/0.3V
Area Per-Bit (mm ² /bit)	0.0246	0.1038	0.0056	0.0056 (RCV)	0.0046
Energy Efficiency (pJ/bit)	1.17	N/A	1.07	0.37 (RCV)	0.29

* 9Gb/s/pin and 8Gb/s/pin for MISR and cell access, respectively

Outline

- Introduction
- Proposed Architecture
- Implementation
- Measurement Results
- Conclusion

Conclusion

- **Digital HBM3 interface is implemented in a 4nm technology**
- **High-Speed, Compact, and Reliable HBM3 interface**
 - Slim bit-slice architecture with stacked I/O
 - Implemented compact PHY size
 - Resistor-tuned offset-calibration with coarse VREF
 - Improved read UI by 16.7% and achieved 300mV operation
 - Digital delay compensation with the accurate delay sensor
 - Achieved voltage noise tolerance up to 220mV
 - In-situ margin-detection
 - Eliminated periodic read training
- **Achieved best PPA with improved reliability**
 - 9.0Gb/s/pin (1.15TB/s/cube), 0.29pJ/bit and 0.0046mm²/bit

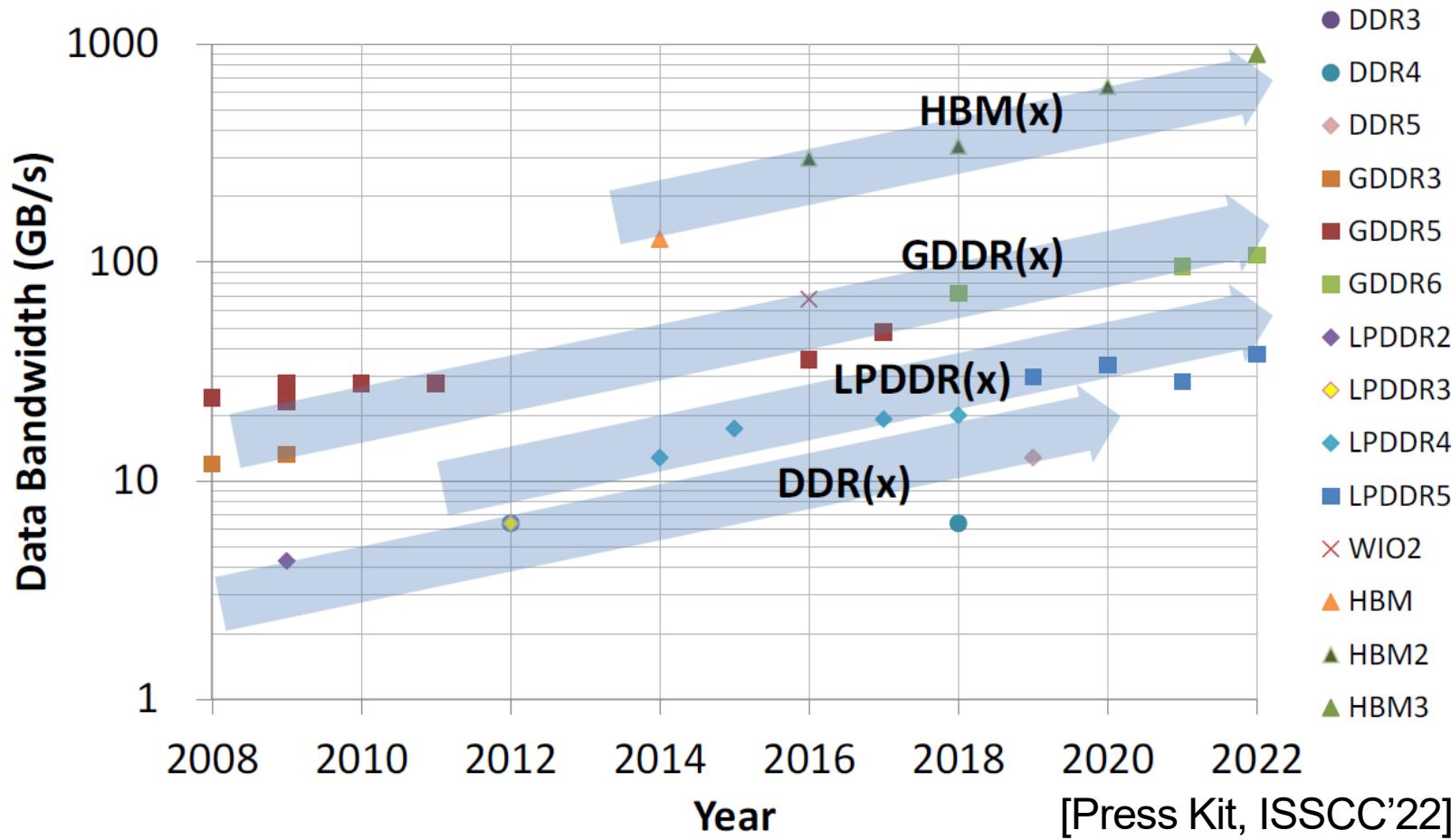
PPA: Power, Performance, and Area

28.5: A 900 μ W, 1–4GHz Input-Jitter-Filtering Digital-PLL-Based 25%-Duty-Cycle Quadrature-Clock Generator for Ultra-Low-Power Clock Distribution in High-Speed DRAM Interfaces

Yuhwan Shin*, Yongwoo Jo*, Juyeop Kim,
Junseok Lee, Jongwha Kim, and Jaehyouk Choi

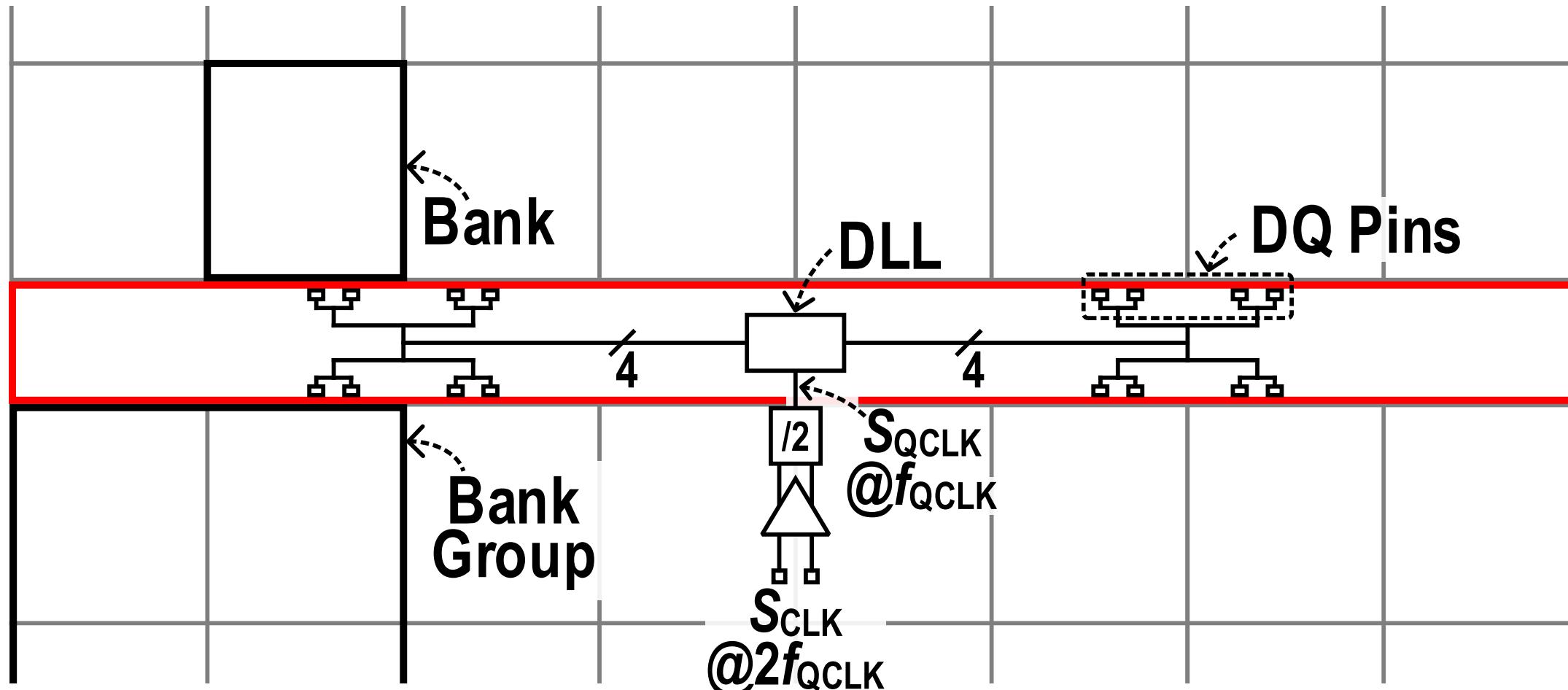
KAIST, Daejeon, Korea
(*Equally-Credited Authors)

Demand on High-Speed DRAM Interface



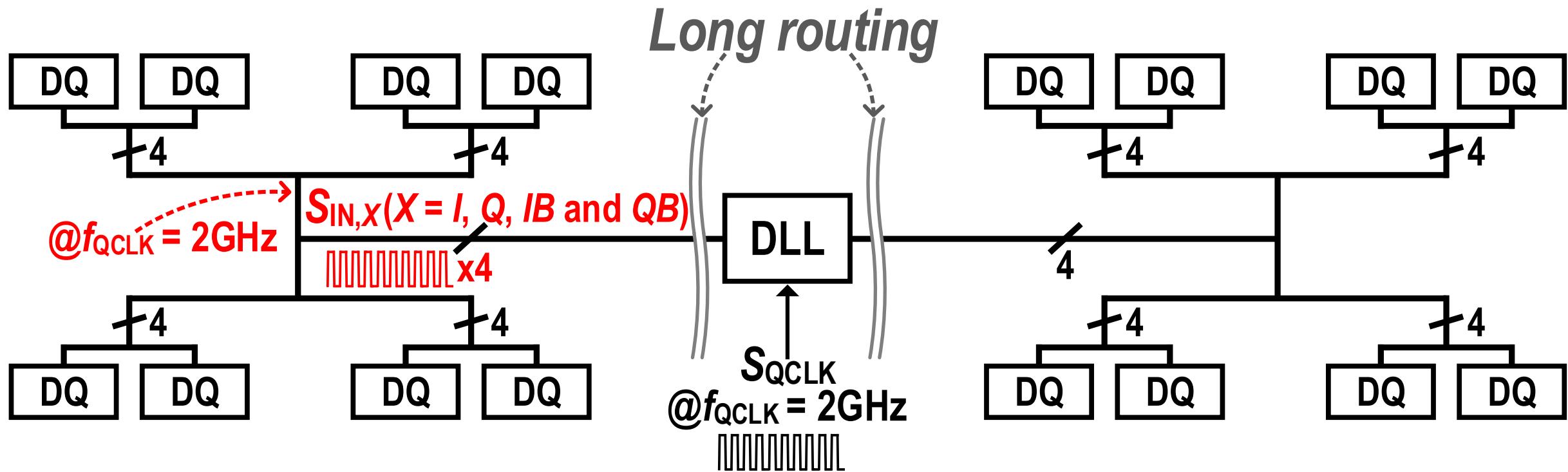
- ❖ Continuous increase in the data bandwidth of DRAM interfaces to accommodate the ever increasing data traffic in various applications

Clock Distribution Scheme on DRAM Chip



- ❖ DRAM interfaces internally using quadrature clocks at a quarter-rate frequency, f_{QCLK}
- ❖ DLL at the middle of peripheral distributes quad. clocks to many DQs across the chip

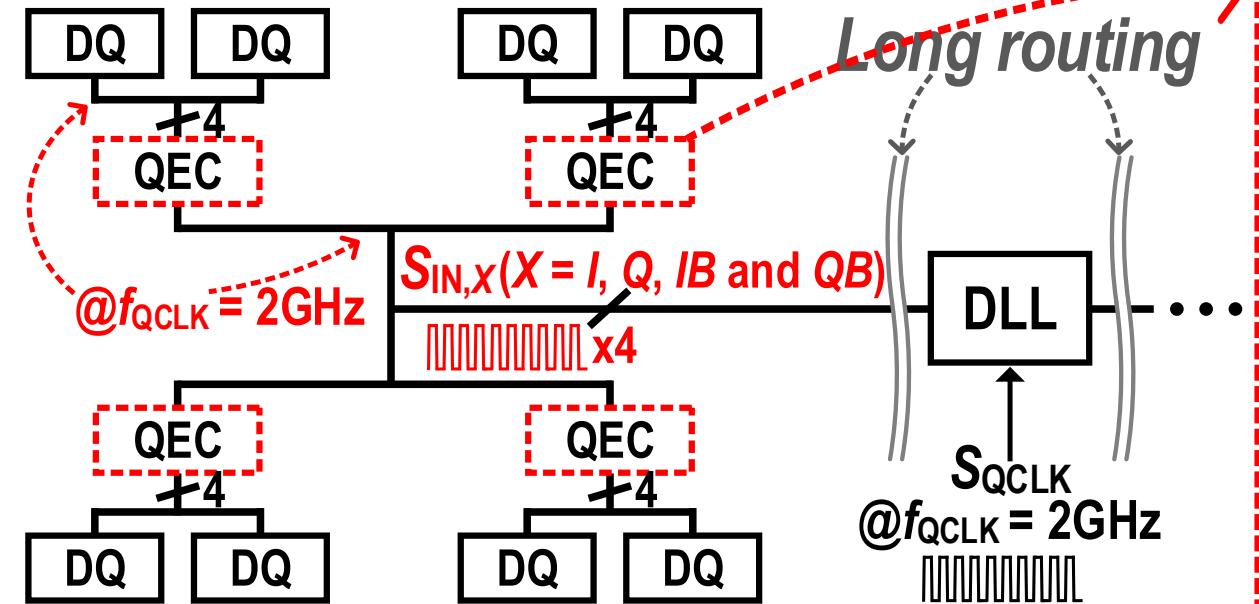
Problems of Conventional Clock Distribution



- ⌚ Large power consumption to distribute quadrature-phase clocks at f_{QCLK} across the chip
- ⌚ Quadrature errors between $S_{IN,X}$ s during long-distance travel

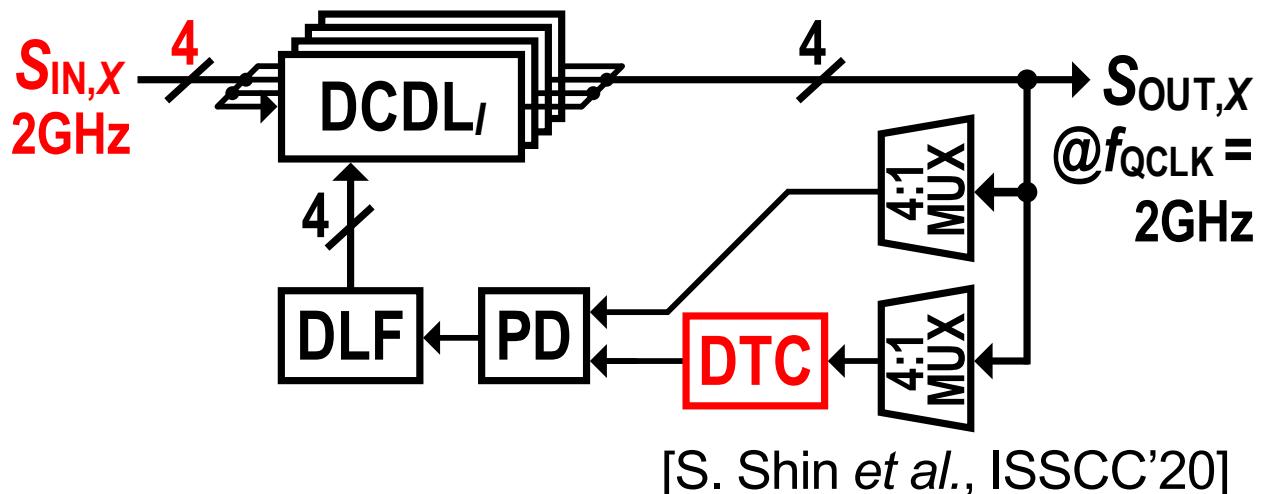
Problems of Conventional Quad.-Error Corrector (QEC)

[Conventional Clock-Distribution]



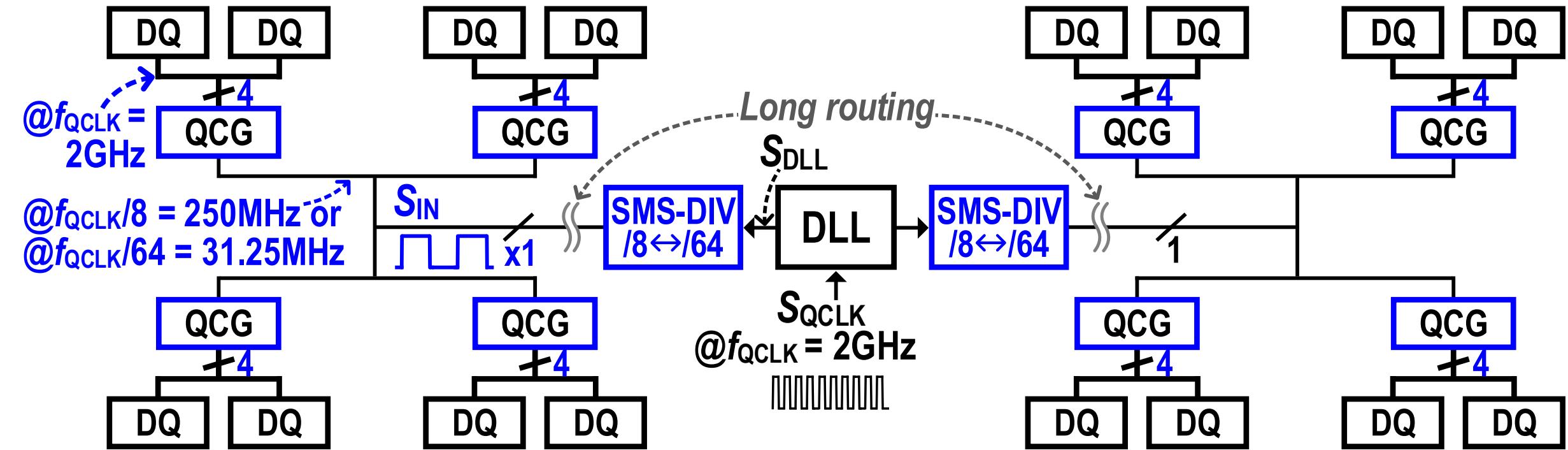
- 😢 Large power to quad.-phase clocks at f_{QCLK}
- 😢 Quad. errors btw $S_{IN,X}$ s during long-dist. travel

[Conventional Quad.-Error Corrector (QEC)]



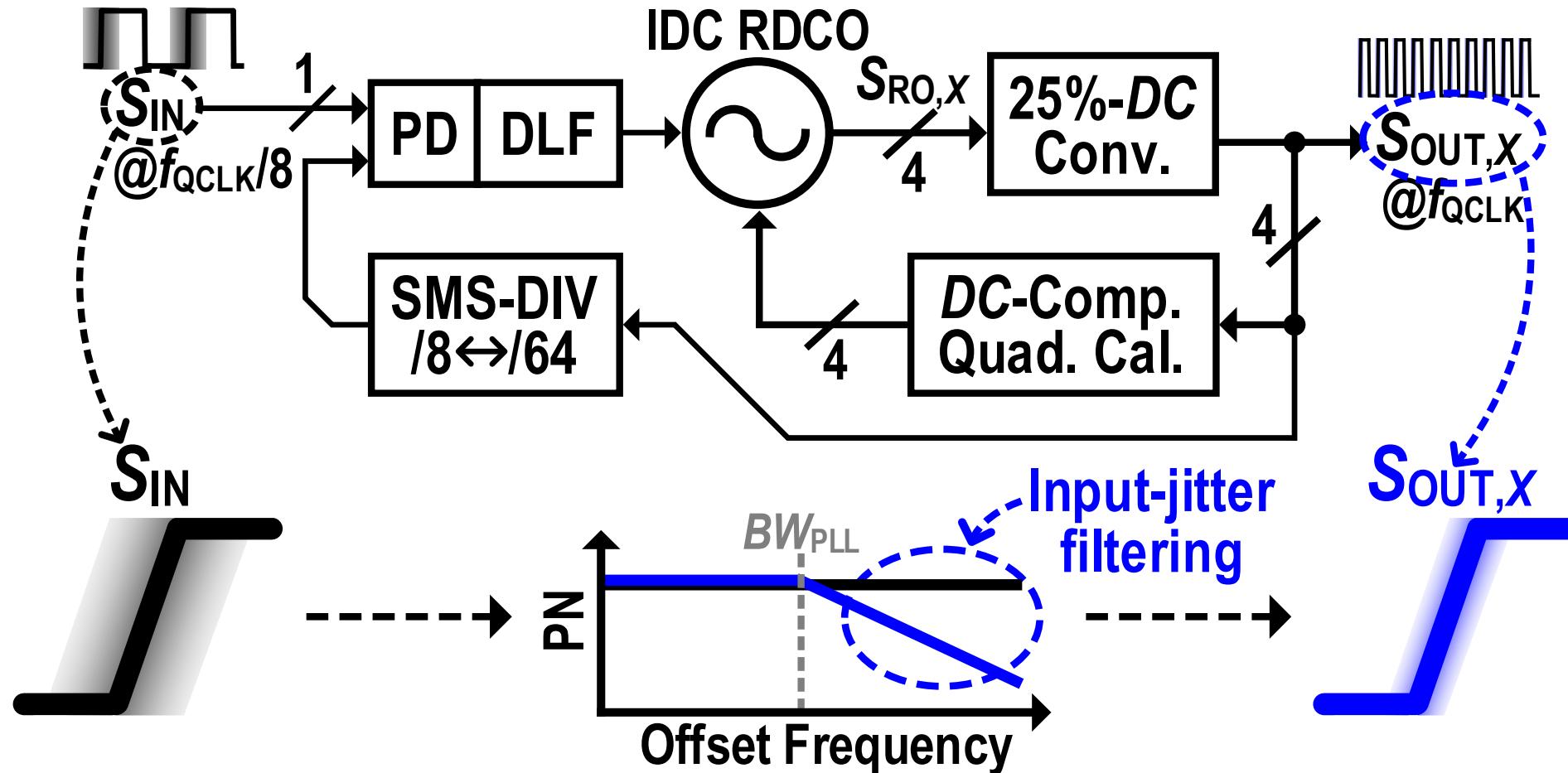
- 😢 Operation $@f_{QCLK} \rightarrow$ Large power consumption
- 😢 DTC-based Quad. Cal. \rightarrow Limited range of f_{QCLK}
- 😢 Additional 25% duty-cycle (DC) conv. required
- 😢 DLL-based \rightarrow No input-jitter filtering

Proposed Low-Power Clock Distribution



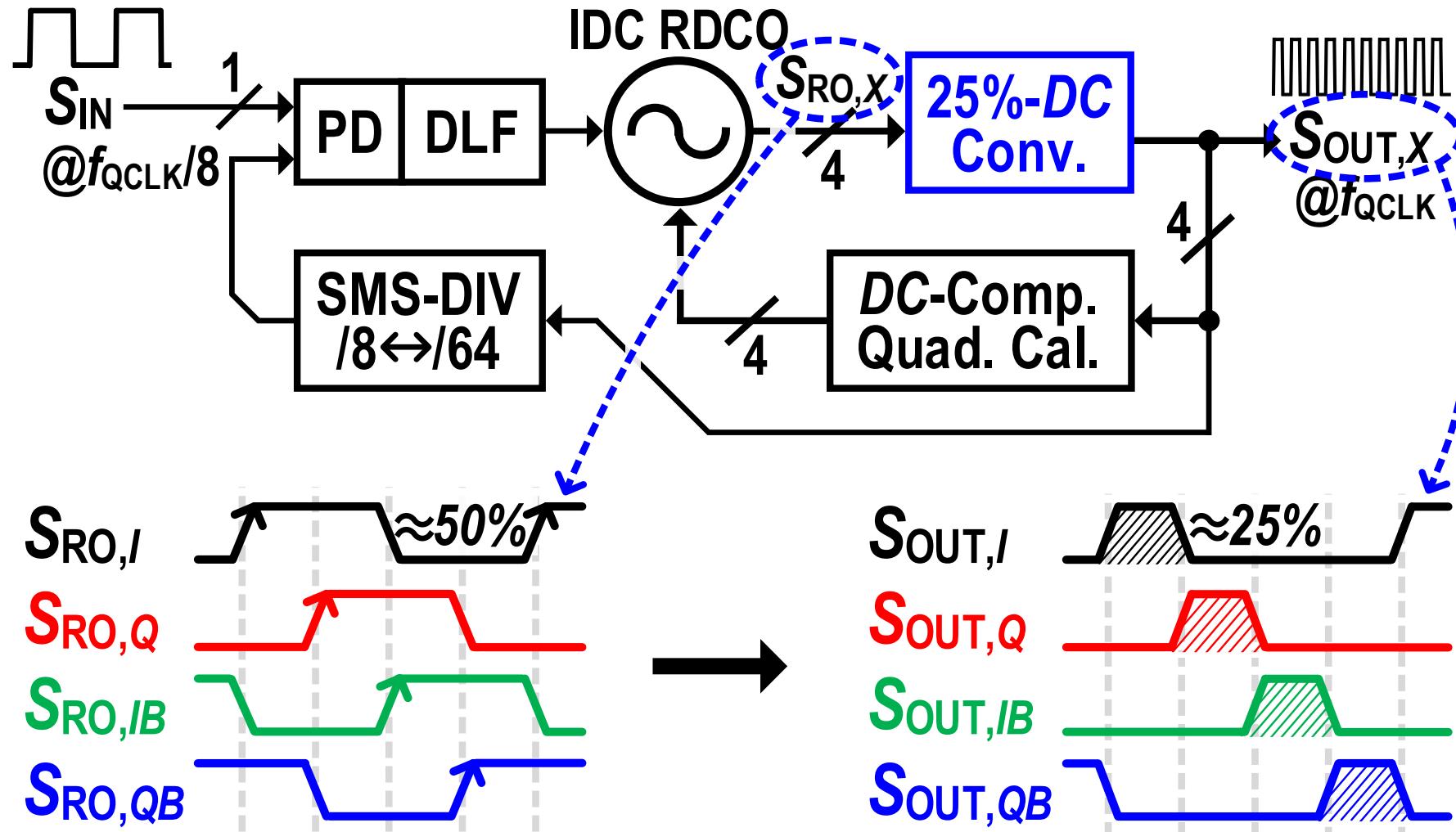
- 😊 Ultra-low power consumption to distribute single-phase clock at a much lower frequency:
 $f_{QCLK}/8$ for *ACTIVE* mode and $f_{QCLK}/64$ for *IDLE* mode
- 😊 Immediate switch. btw. *ACTIVE* & *IDLE* by synchronous mode-switching divider (SMS-DIV)

Proposed Quad.-Clock Generator (QCG): Advantage I



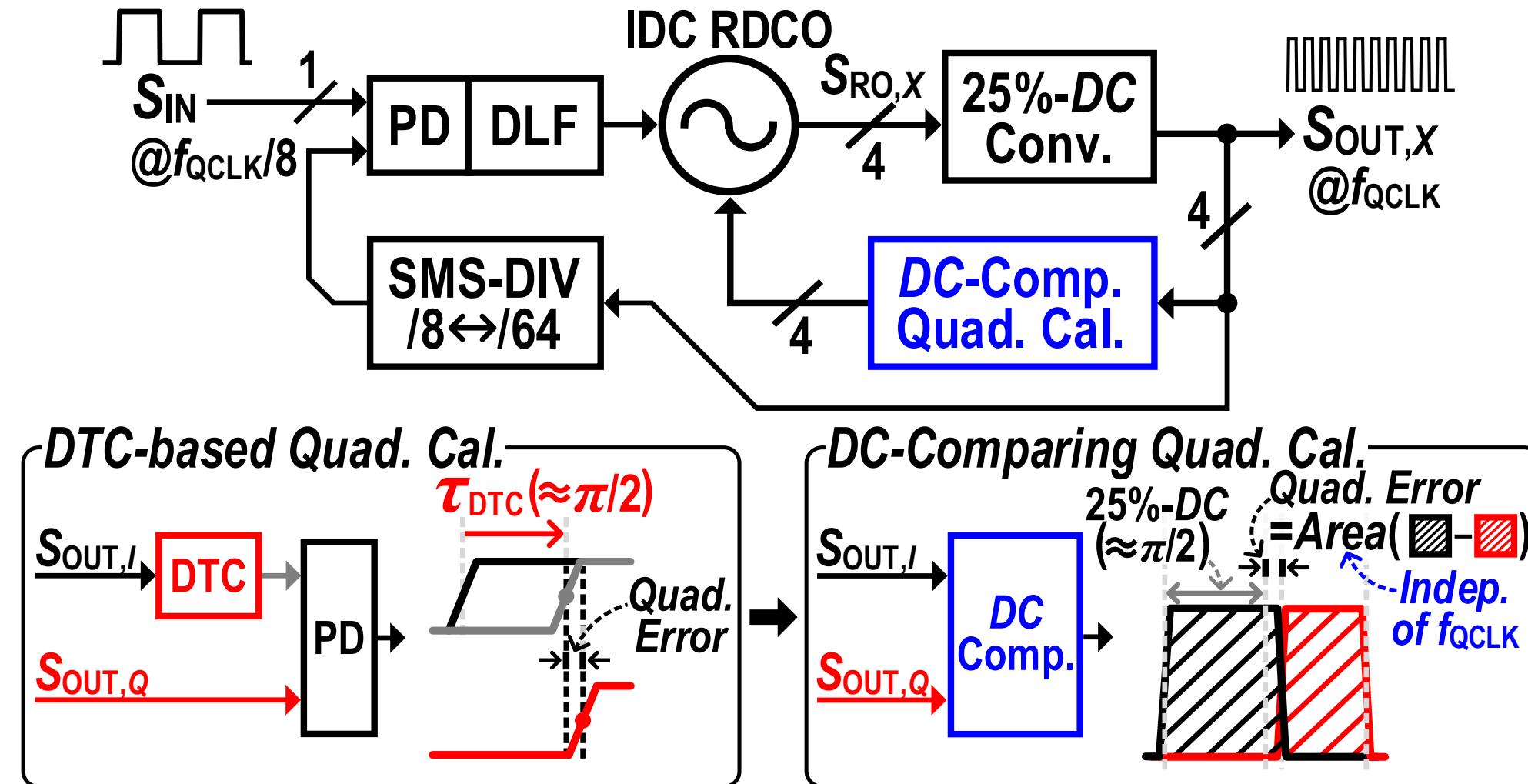
- ☺ Operation $@f_{QCLK}/8 \rightarrow$ much less power consumption of loop-building blocks
- ☺ DPLL-based Architecture \rightarrow dramatical jitter filtering of input clock

Proposed Quad.-Clock Generator (QCG): Advantage II



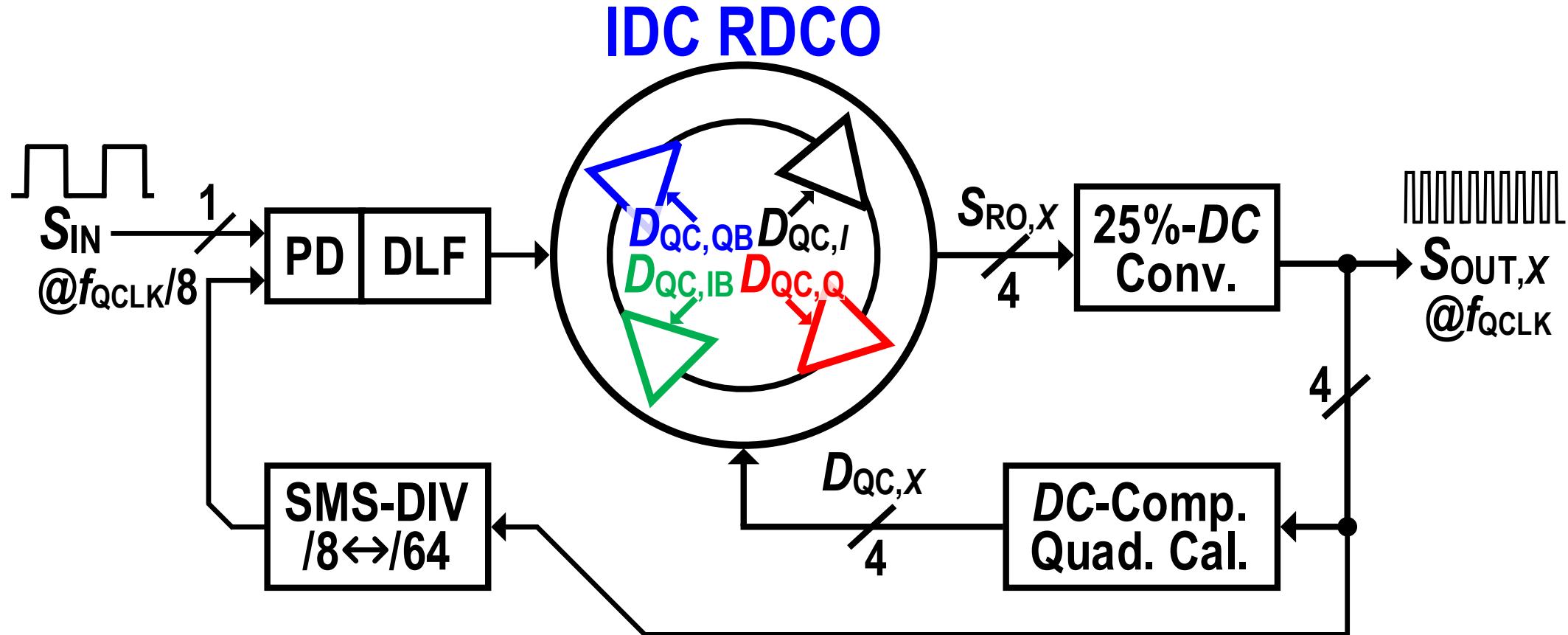
😊 Natural 25% Duty-Cycle (DC) with minimized DQ skew

Proposed Quad.-Clock Generator (QCG): Advantage III



😊 DC-Comparing Quadrature-Error Calibrator → Wide range of f_{QCLK} (1-4GHz)

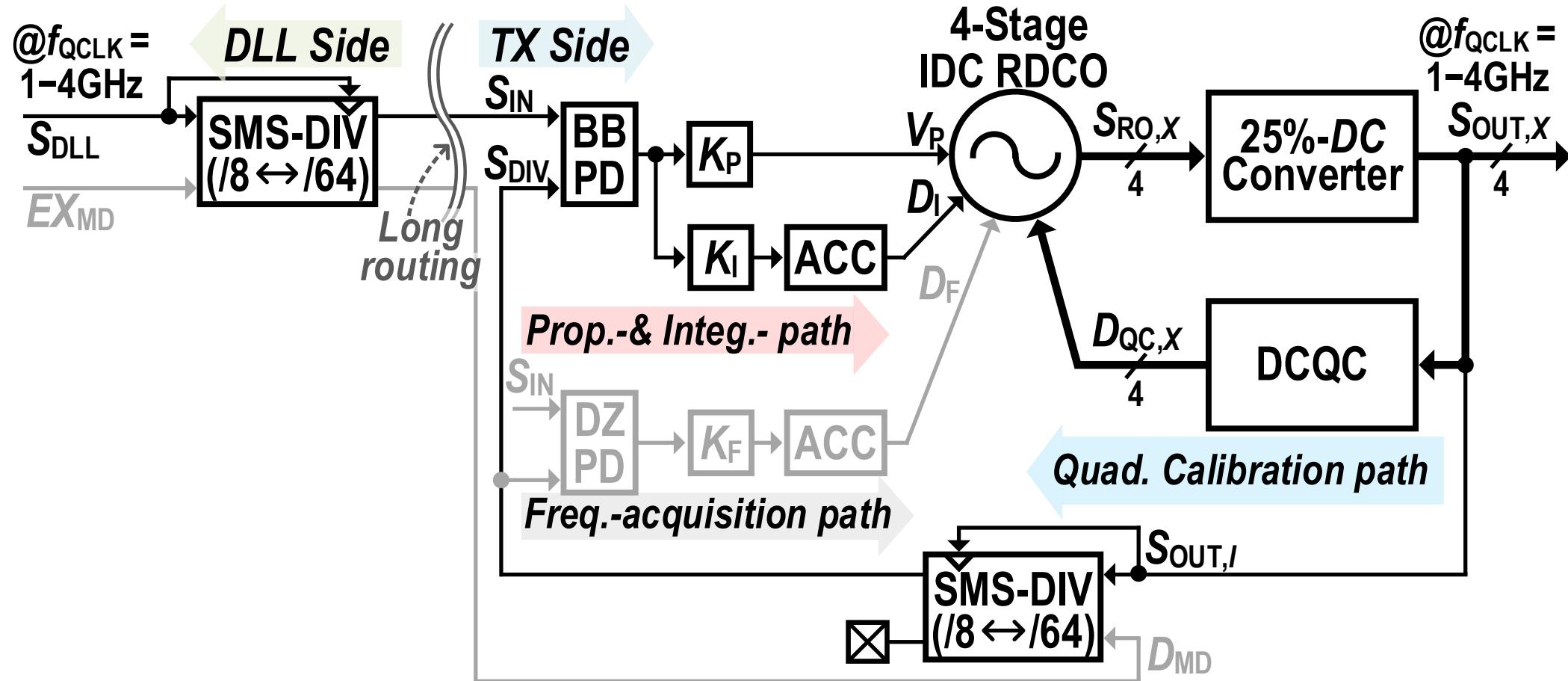
Proposed Quad.-Clock Generator (QCG): Advantage IV



😊 Individual-Delay-Controlling Ring DCO (IDC RDCO)

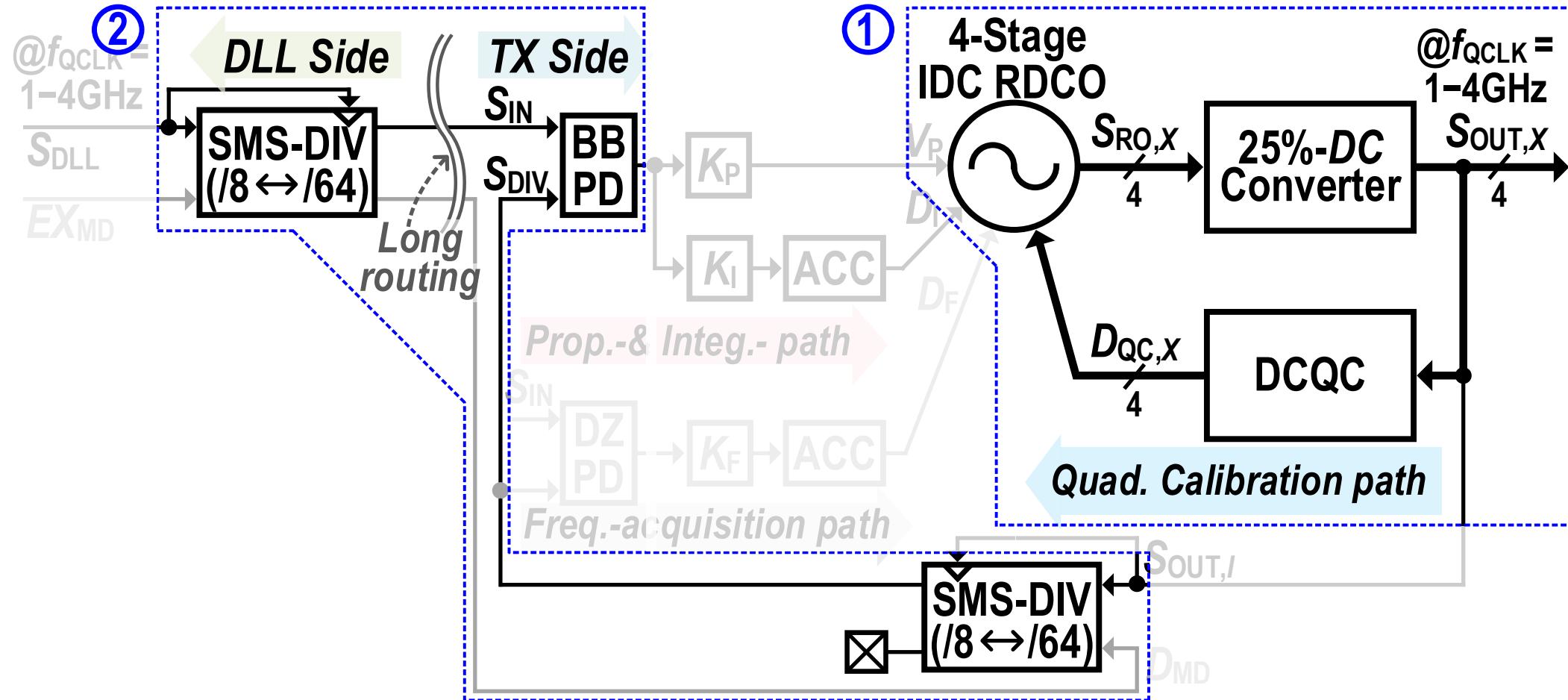
→ No extra delay cell for quadrature-error calibration

Overall Architecture of Proposed QCG



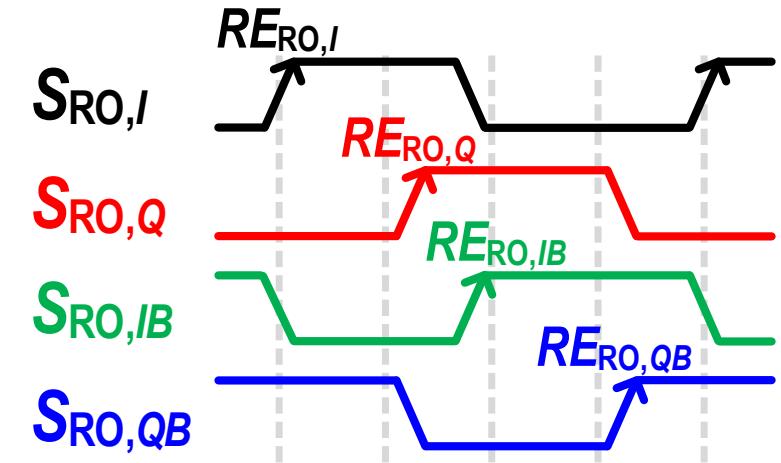
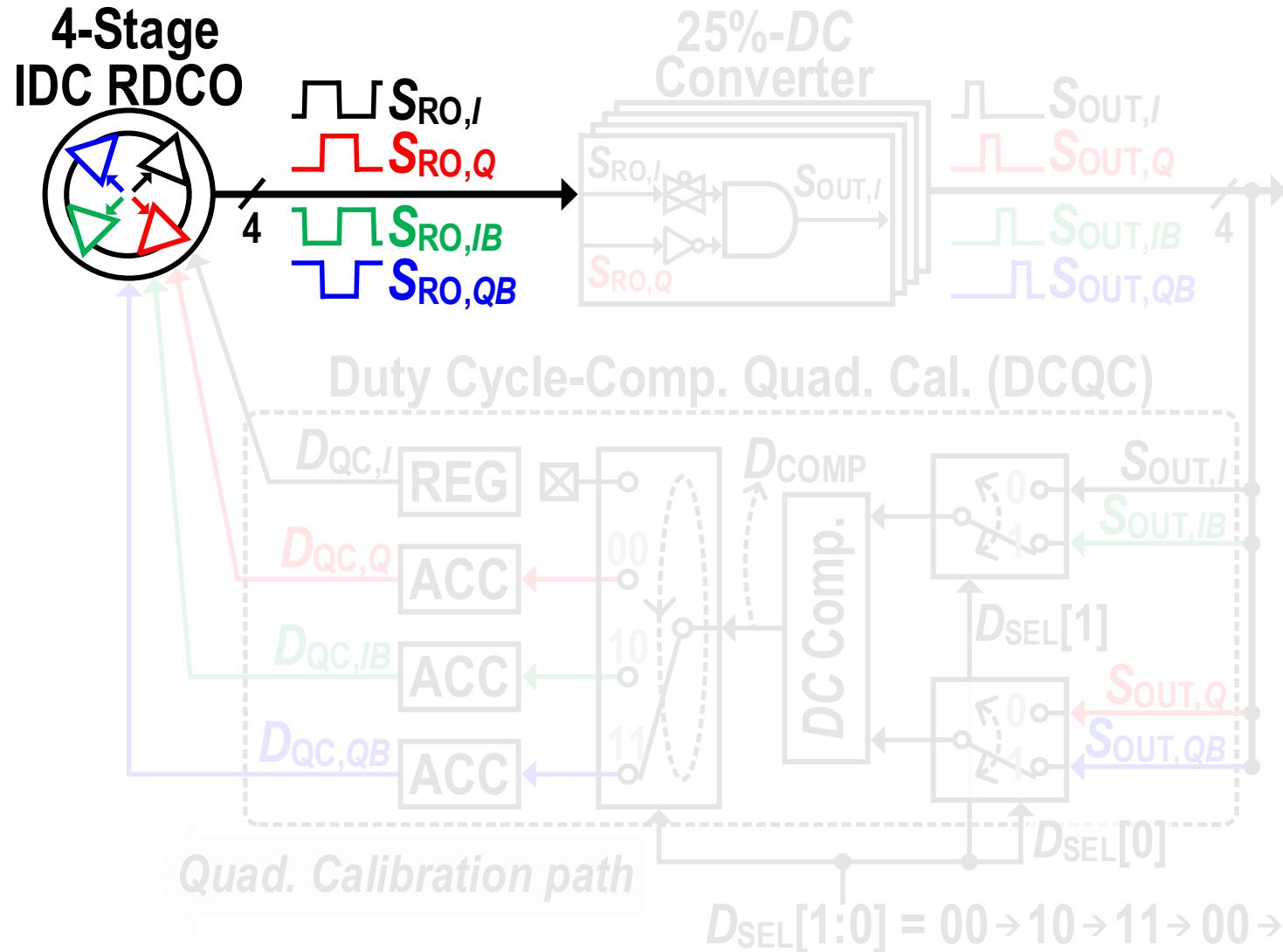
- ❖ Based on type-II DPLL for input-jitter filtering
- ❖ Additional frequency acquisition path for initial locking

Overall Architecture of Proposed QCG

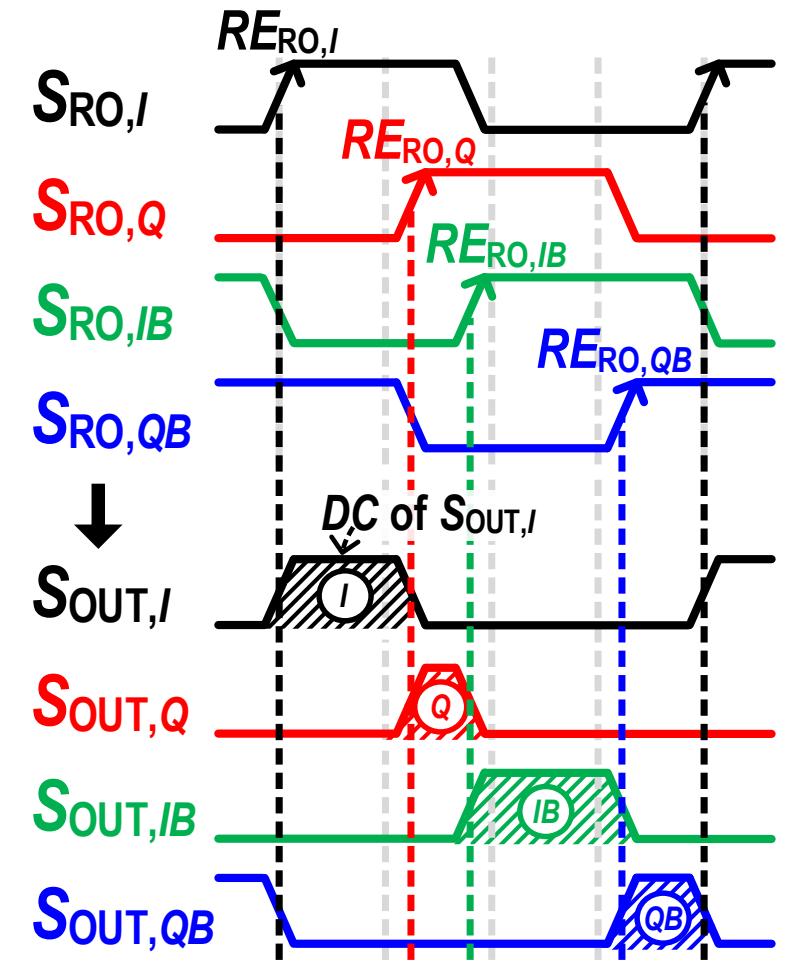
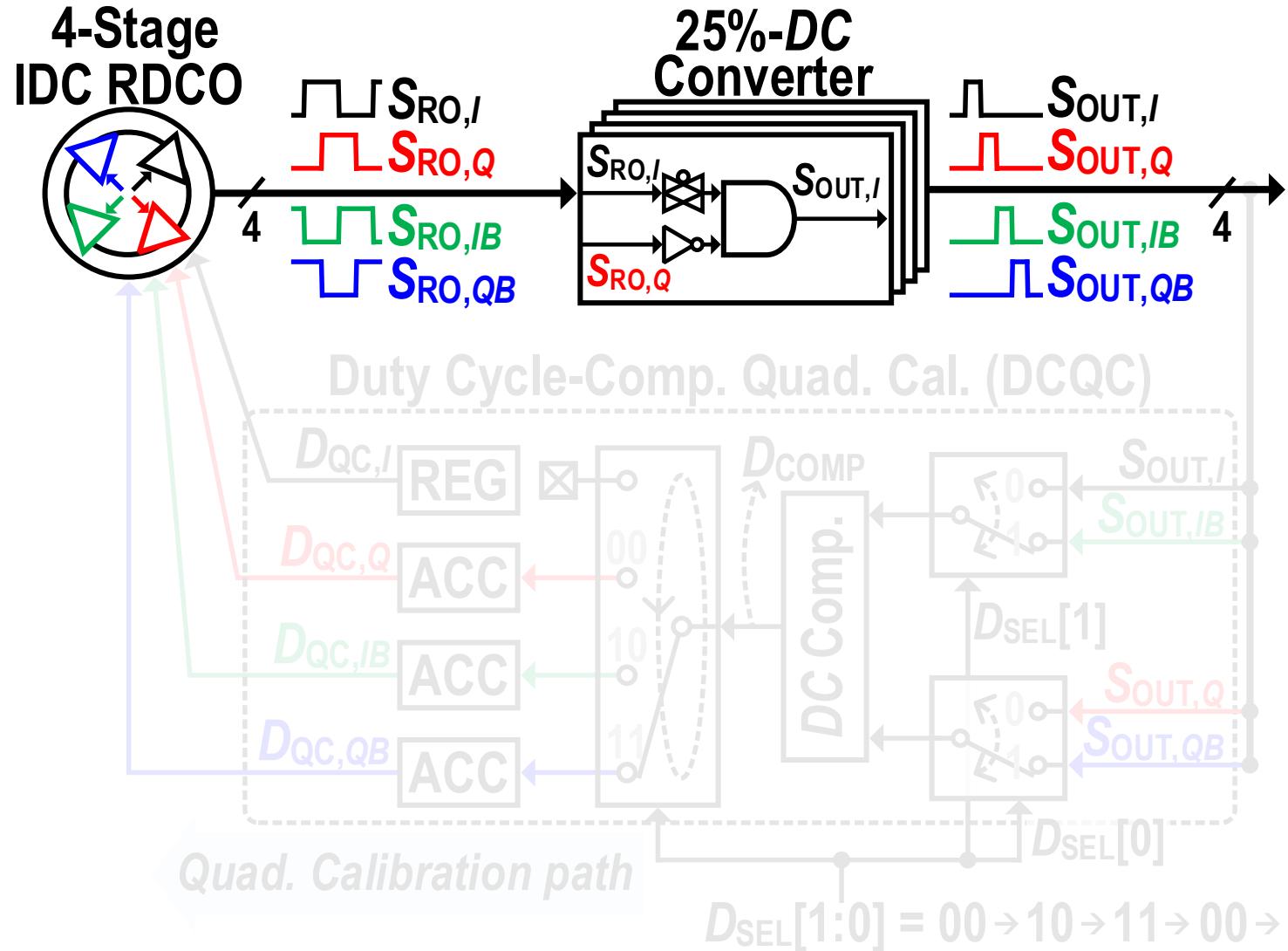


- ① DCQC to correct quadrature errors over a very wide-range of f_{QCLK}
- ② SMS-DIV to enable an immediate return from *IDLE* mode to *ACTIVE* mode

Overall Operation of DCQCC

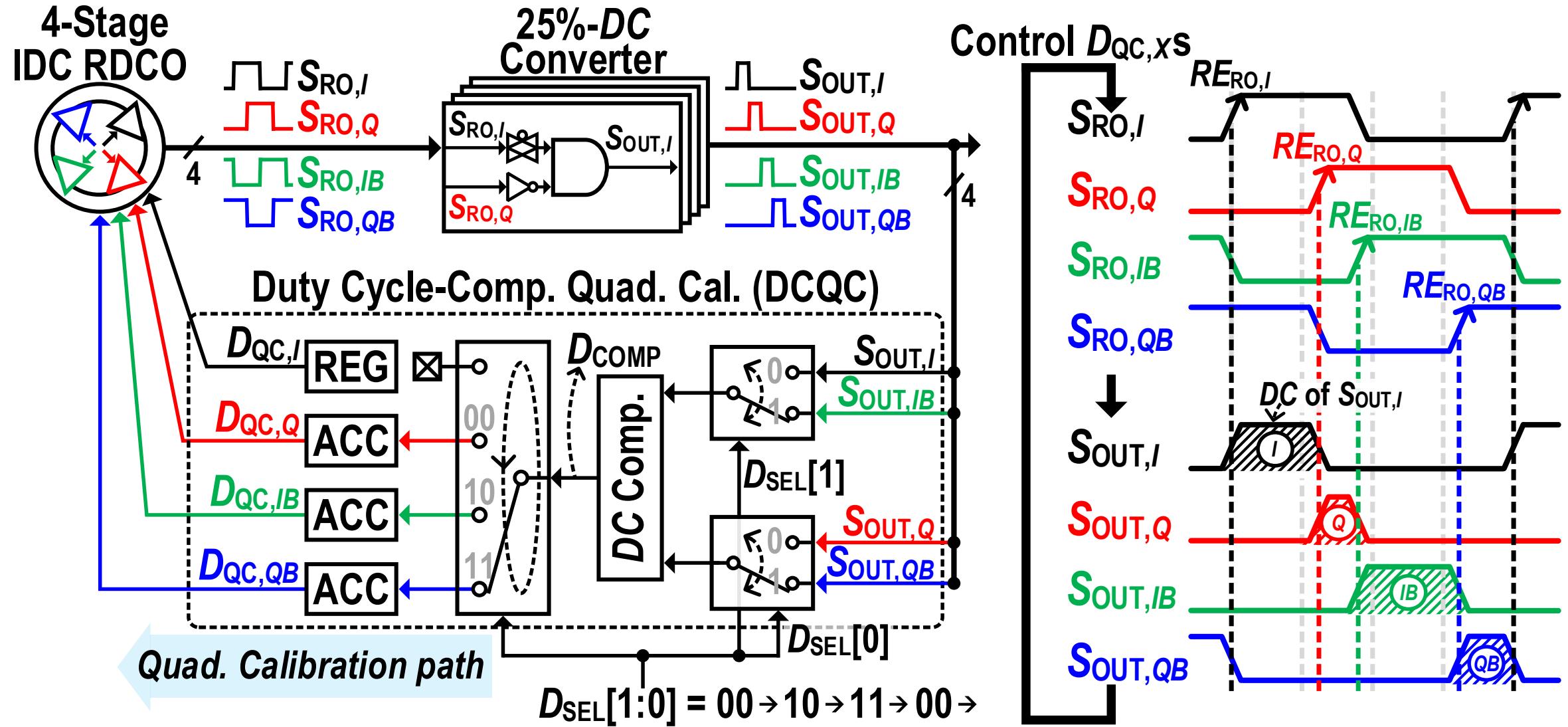


Overall Operation of DCQCC

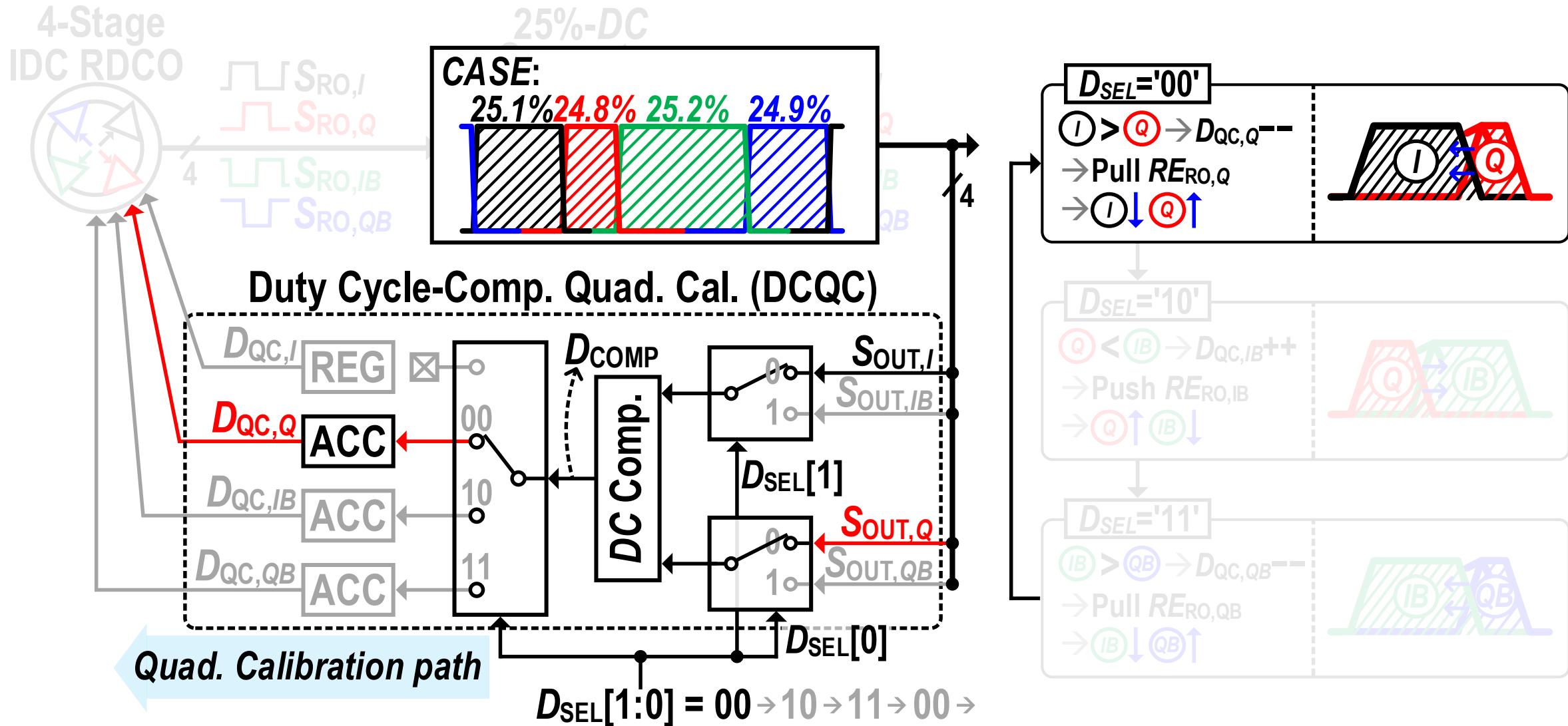


28.5: A 900 μ W, 1–4GHz Input-Jitter-Filtering Digital-PLL-Based 25%-Duty-Cycle Quadrature-Clock Generator for Ultra-Low-Power Clock Distribution in High-Speed DRAM Interfaces

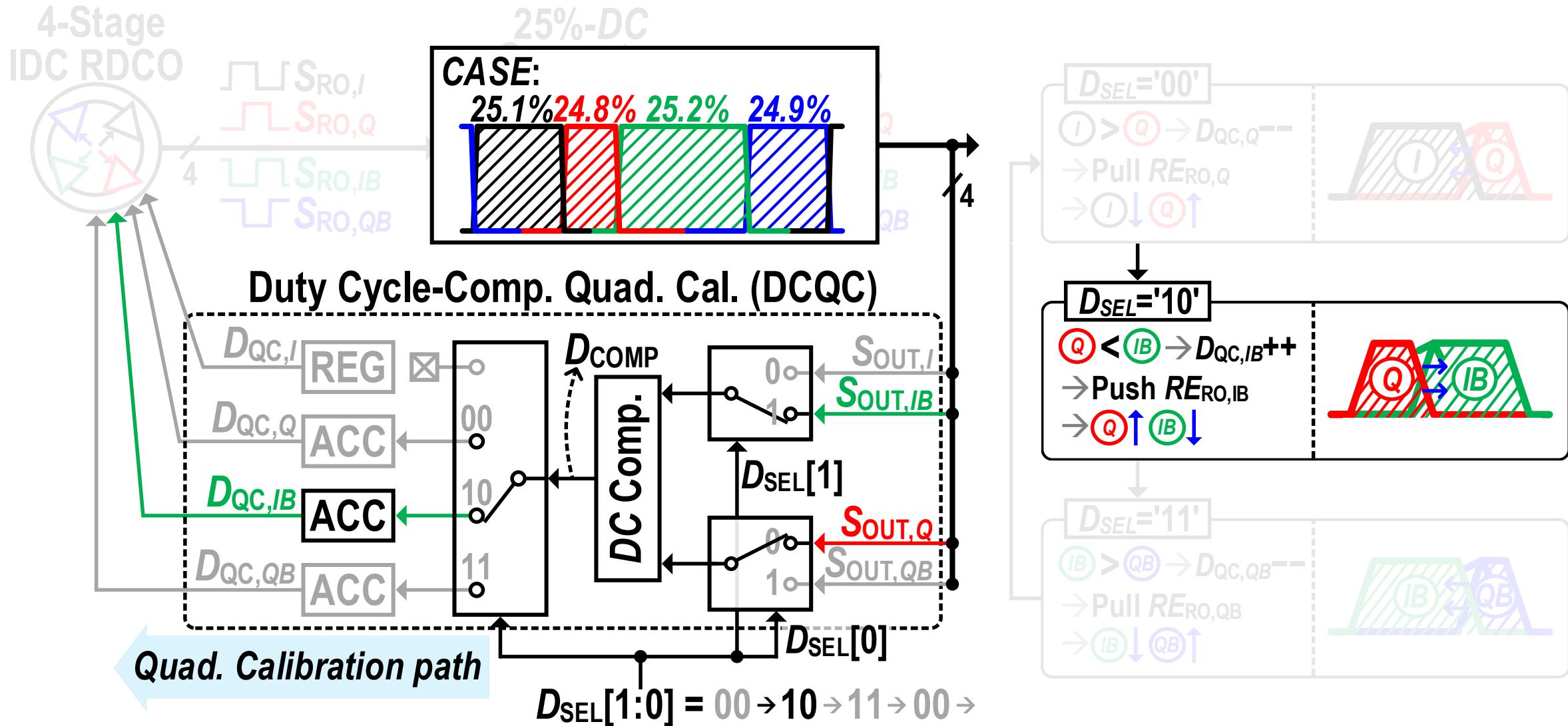
Overall Operation of DCQC



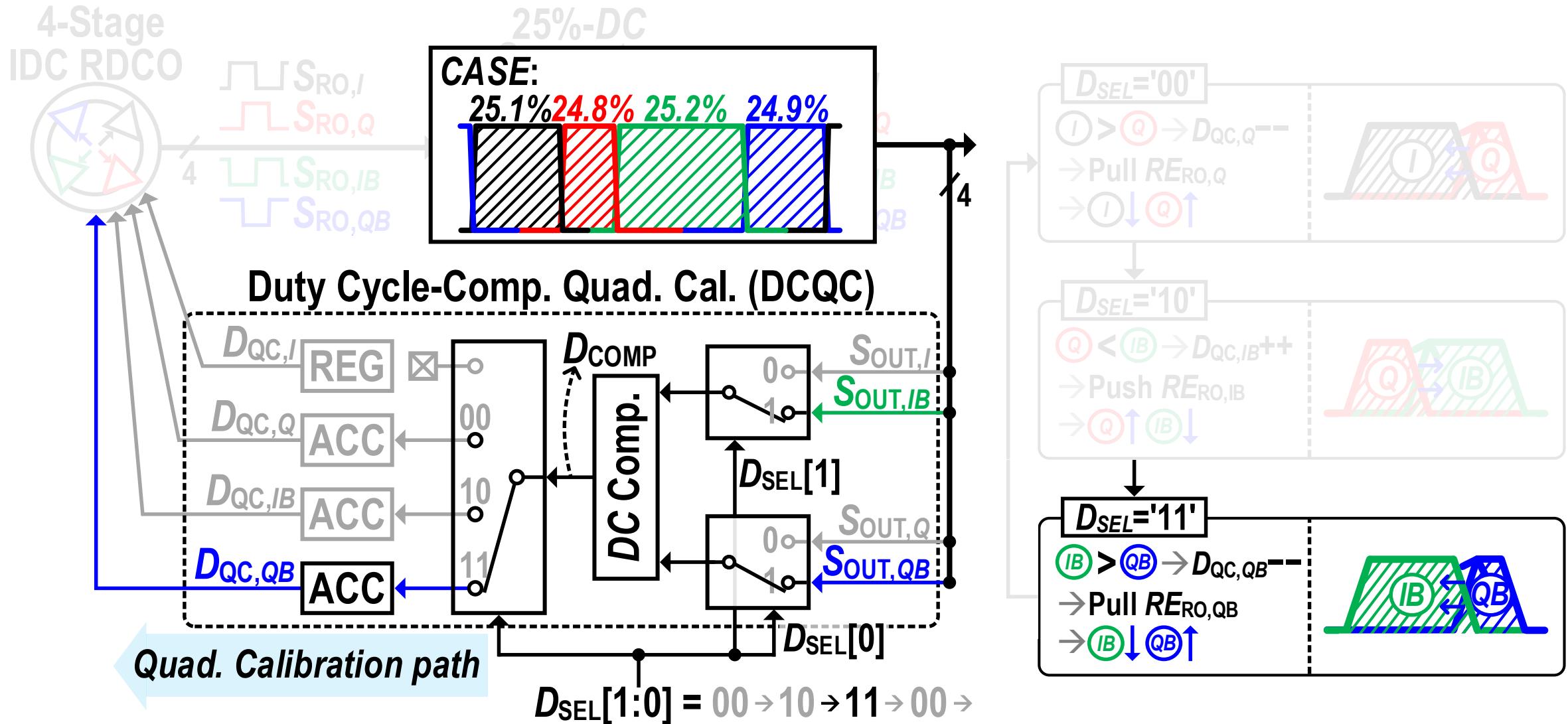
Overall Operation of DCQC



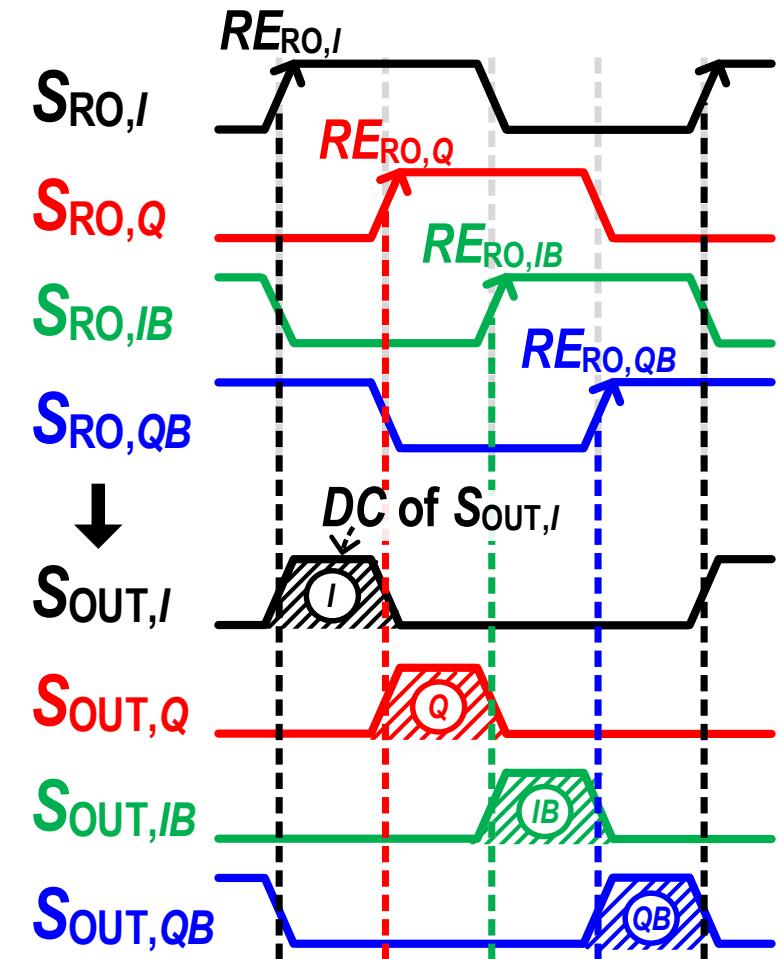
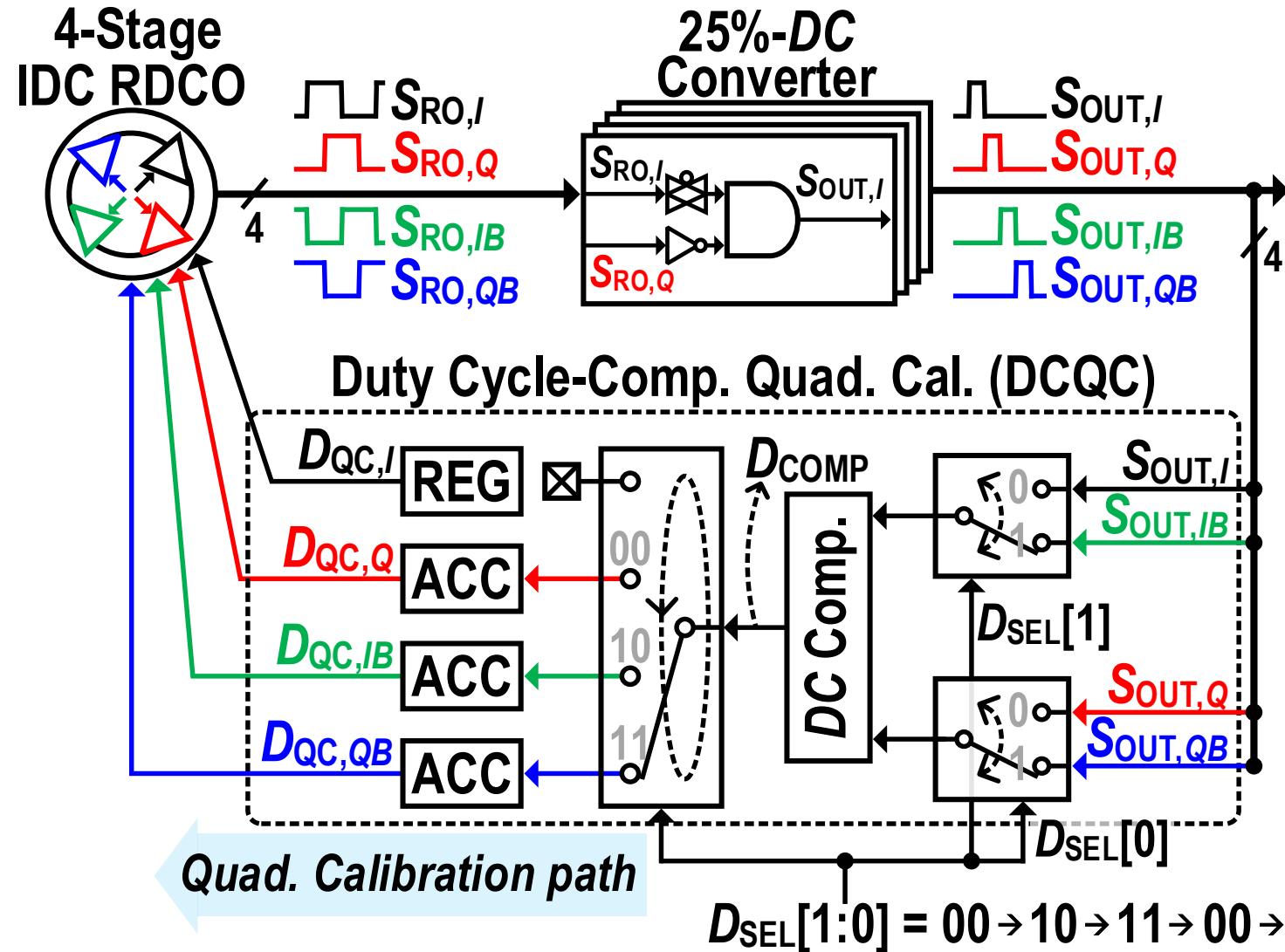
Overall Operation of DCQC



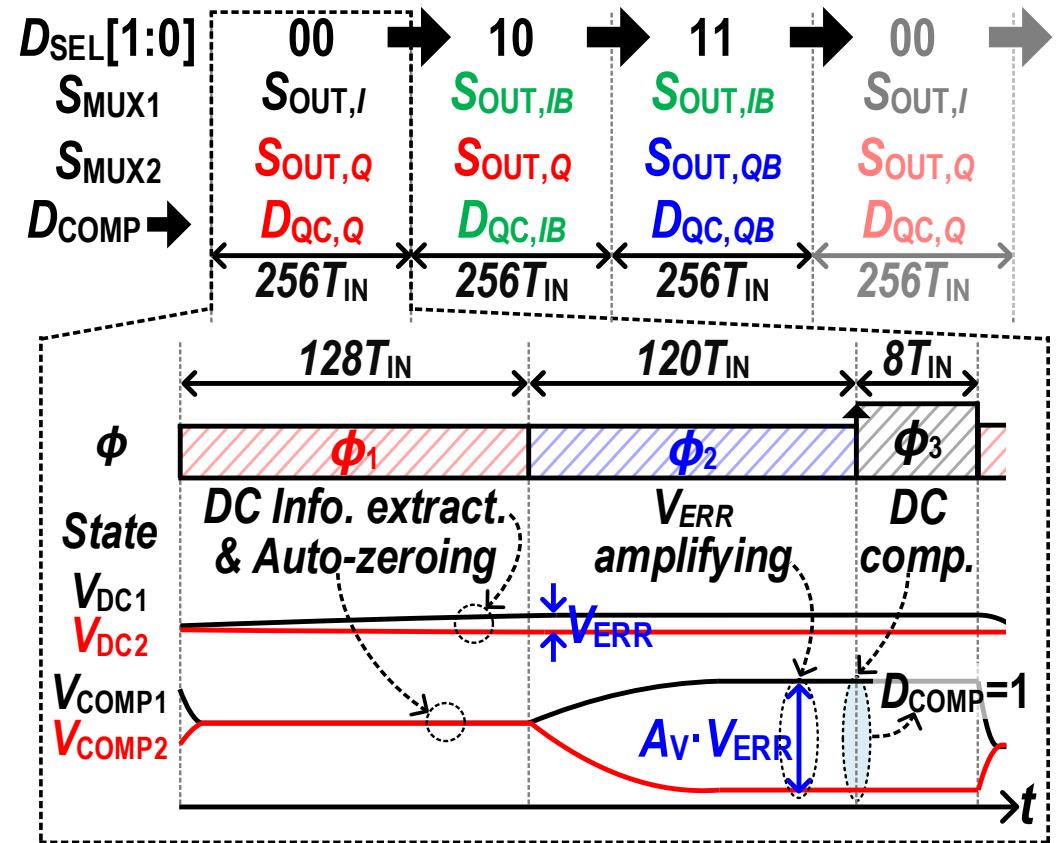
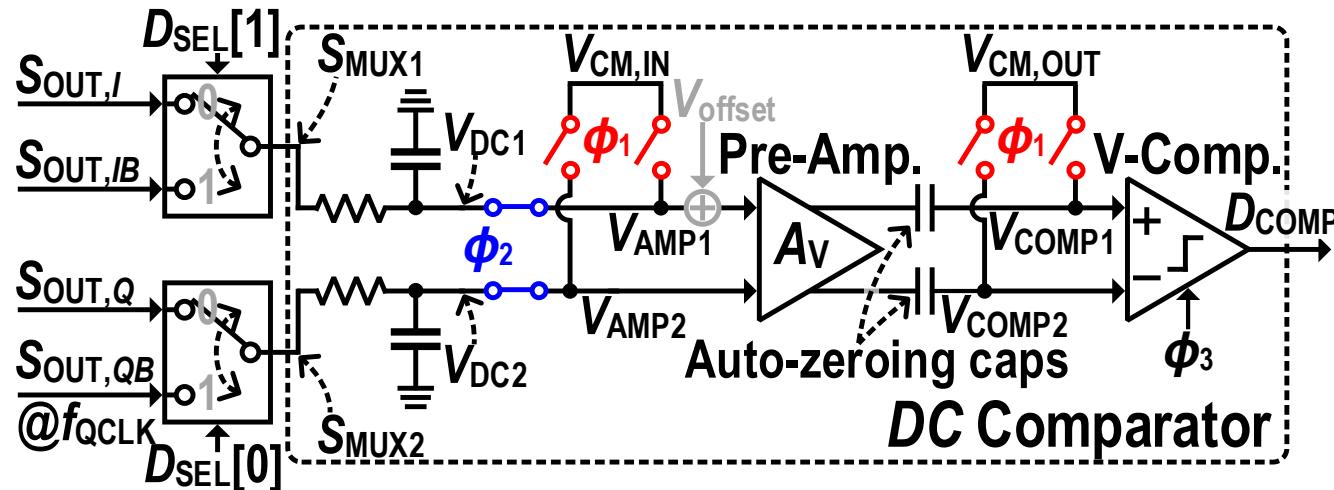
Overall Operation of DCQC



Overall Operation of DCQC

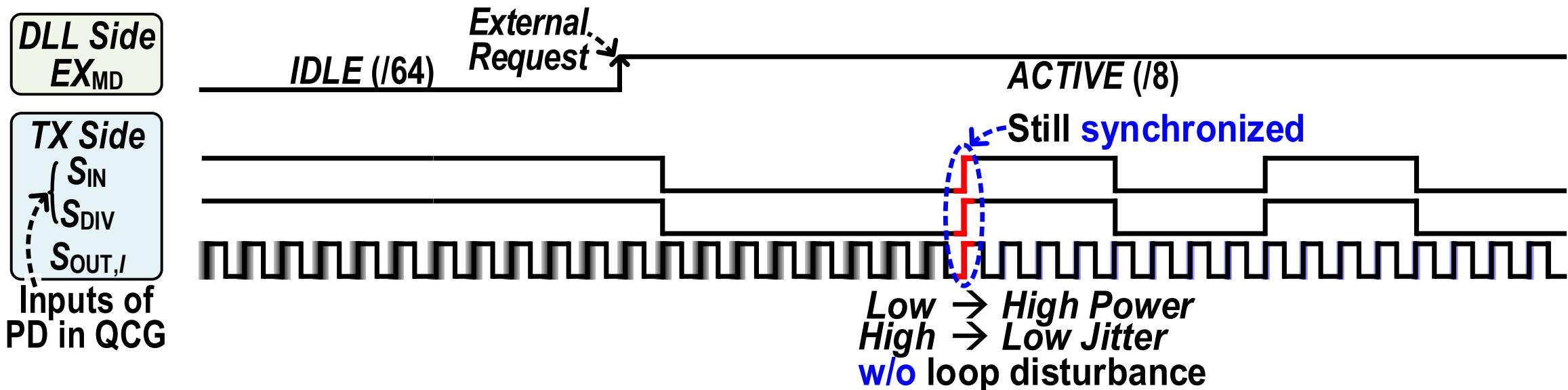
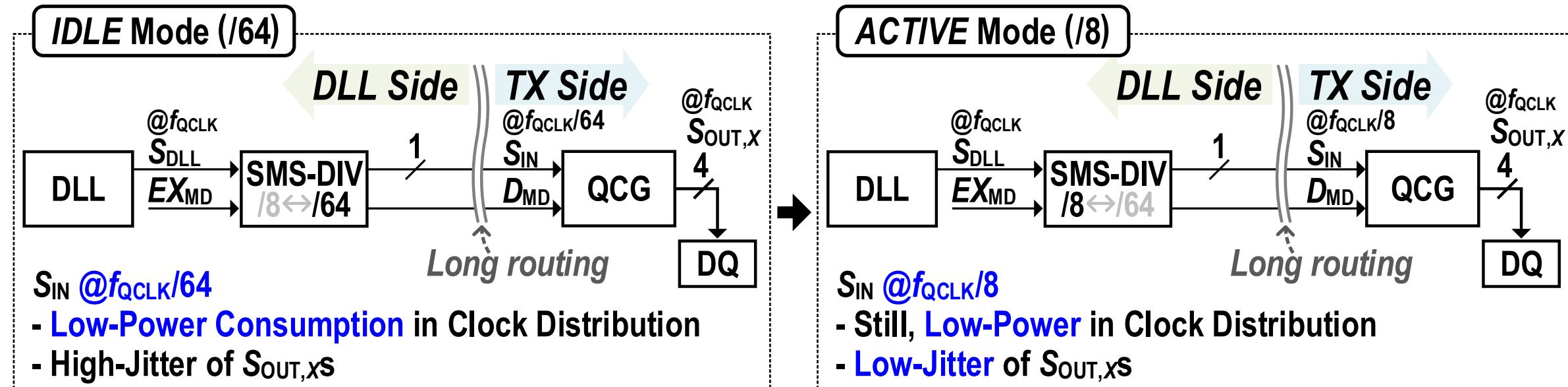


Design of DC Comparator in DCQC

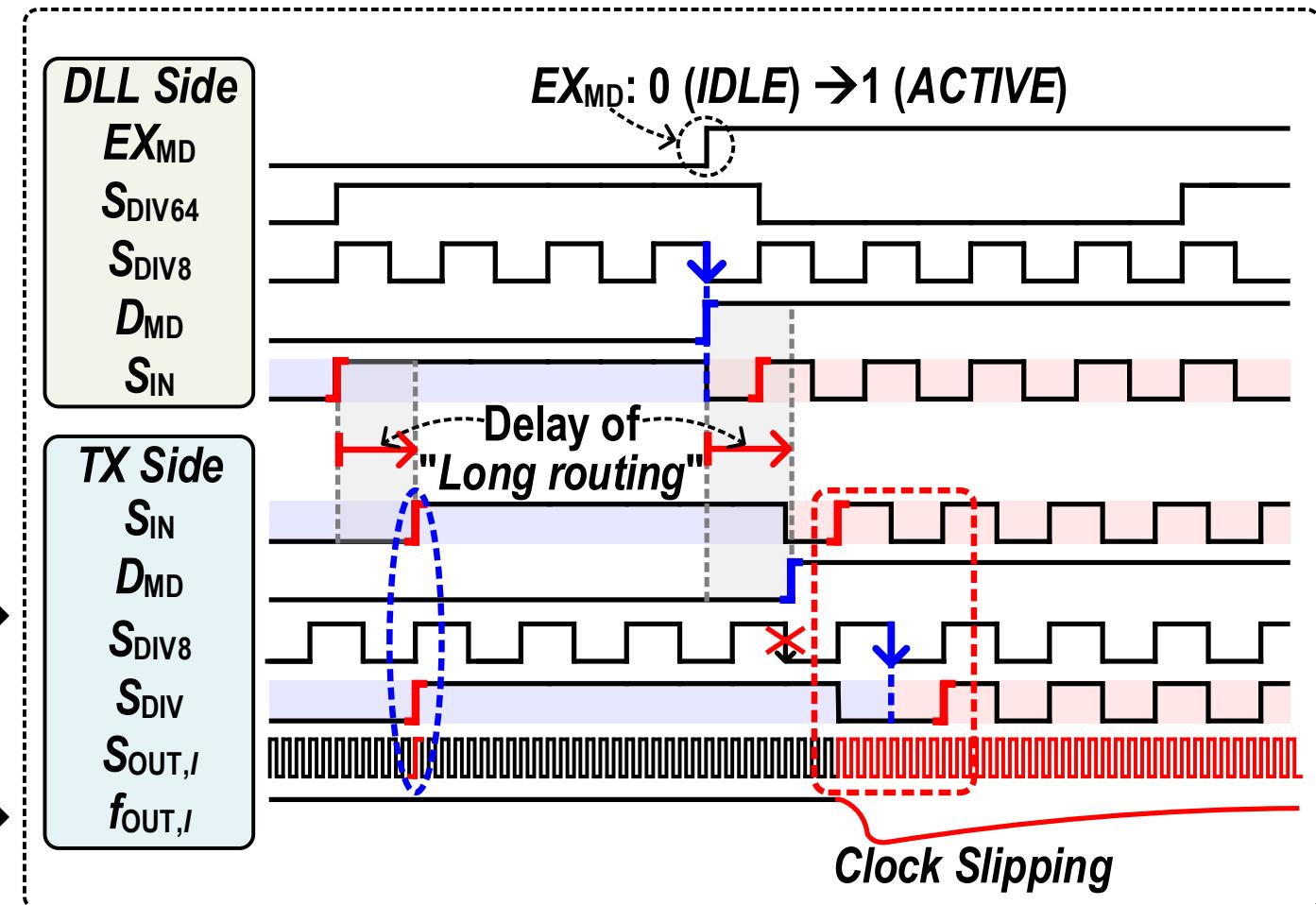
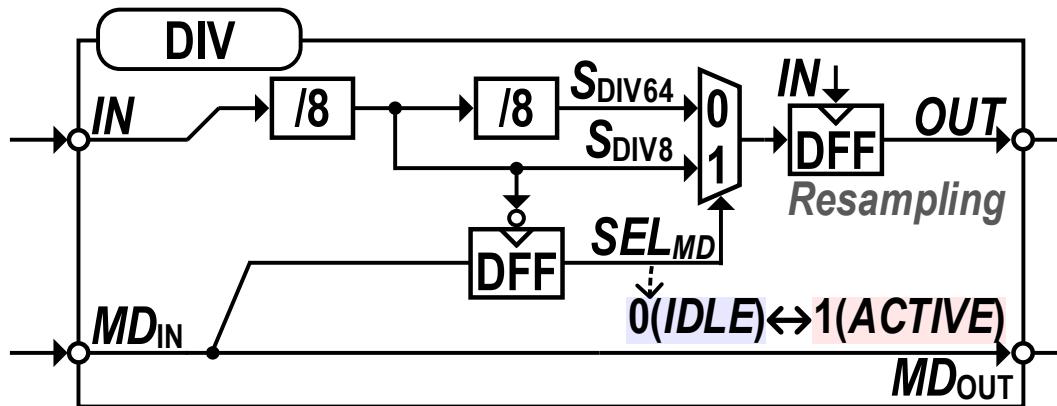
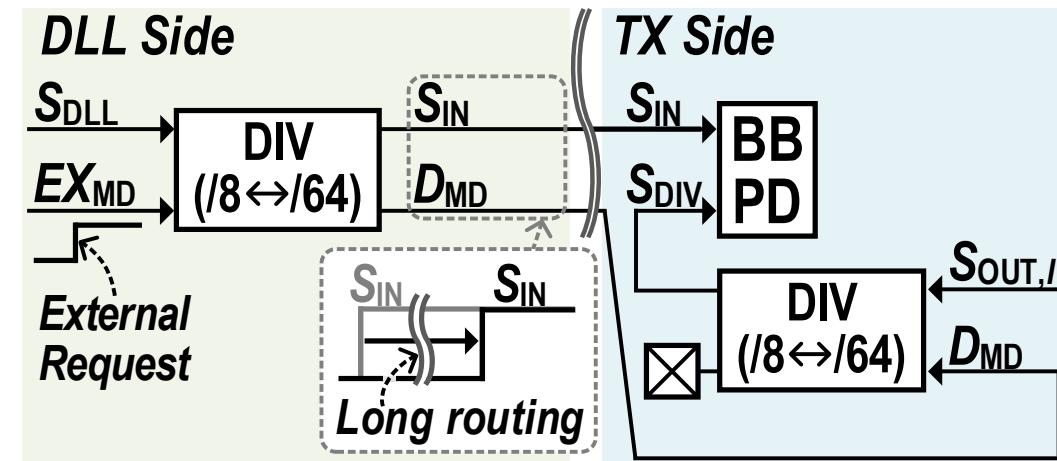


- ❖ Extract DCs → Pre-amplify the difference of voltage → Compare DCs
- ❖ Auto-zeroing technique → Cancel the offset voltage of the DC Comparator

Objective of Sync. Mode-Switching Divider (SMS-DIV)

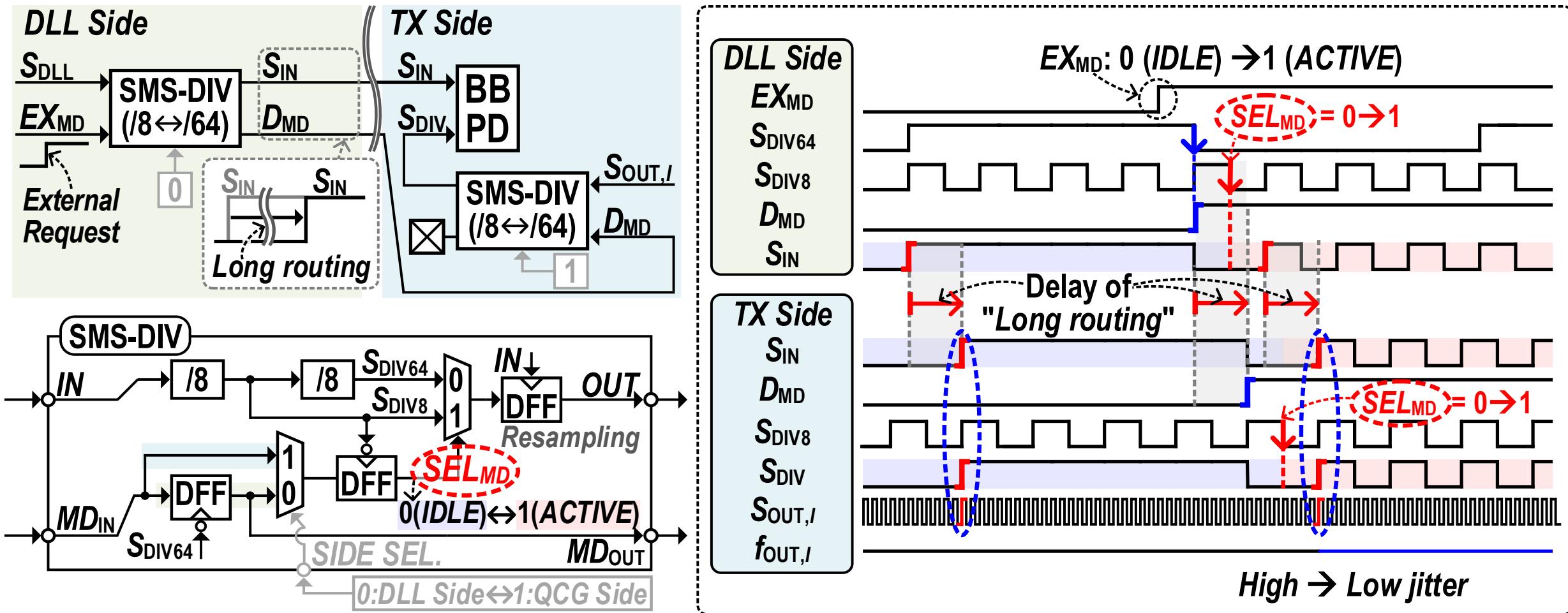


Problems of Mode-Switching w/o SMS-DIV



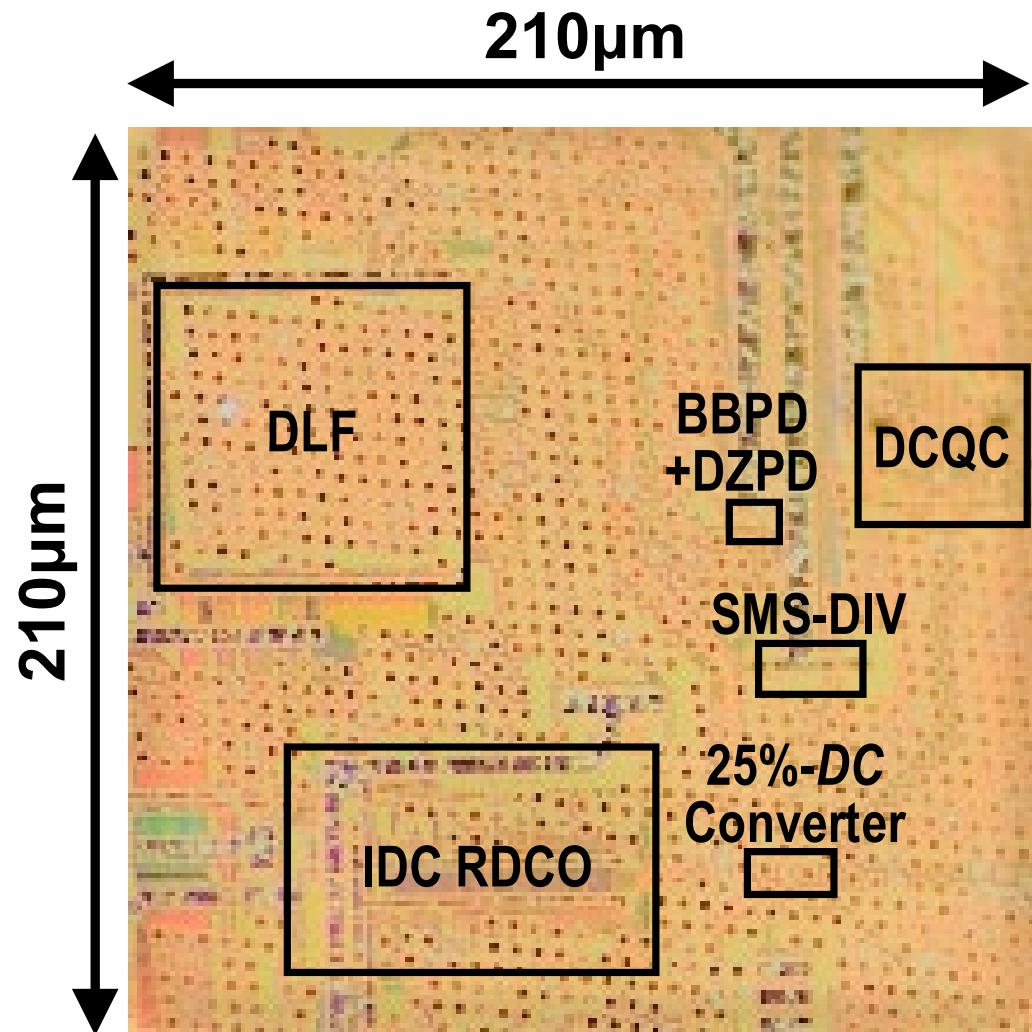
- ❖ Instantly, S_{IN} and S_{DIV} are out of alignment
- The QCG requires **resetting time** to settle **loop disturbance**

Mode-Switching w/i SMS-DIV



- ❖ S_{IN} and S_{DIV} can stay aligned despite the mode switching
- The QCG can maintain a lock **without any loop disturbance**

Die Photograph

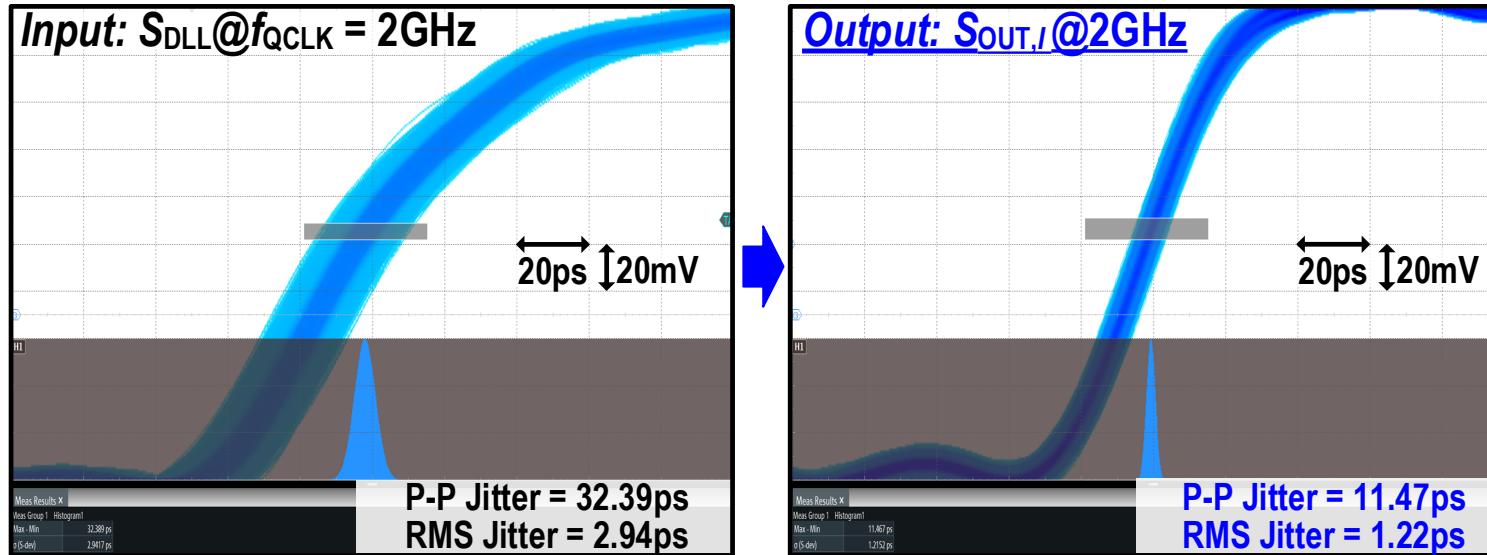


❖ 40nm CMOS

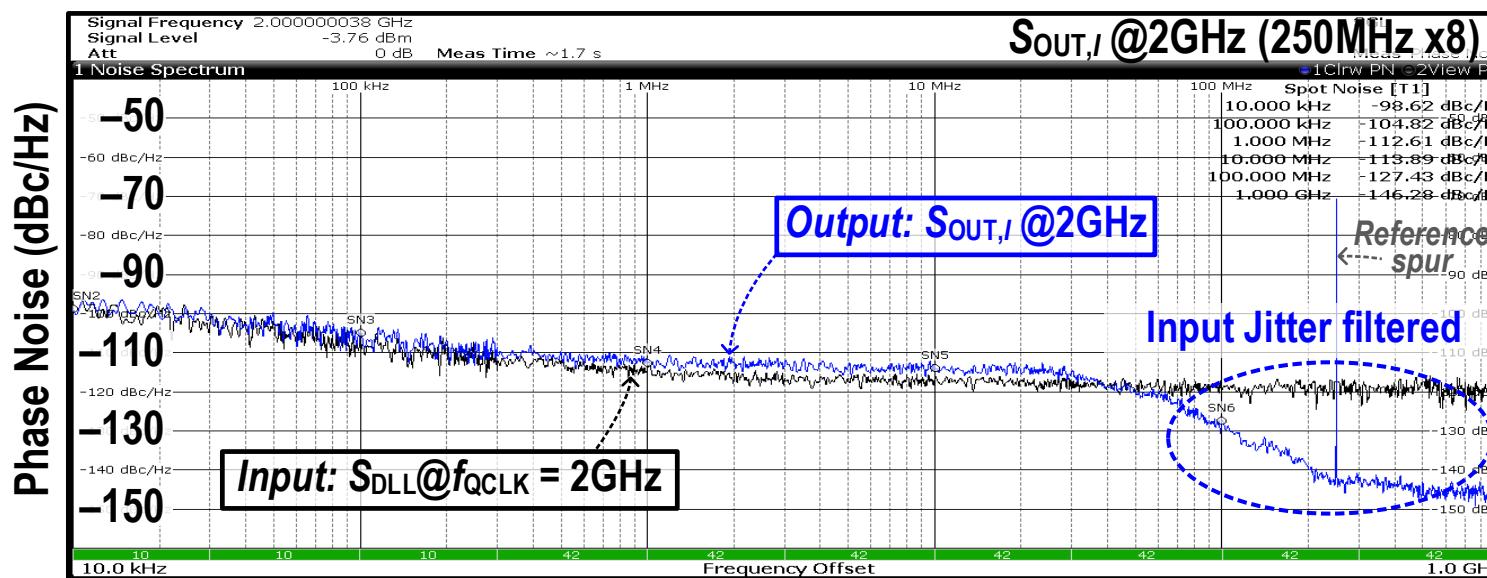
Power Consump. (mW) in ACTIVE mode	
IDC RDCO	0.5
DLF	0.2
25%-DC Converter	0.05
DCQC (DC Comparator only)	0.1 (0.05)
SMS-DIV & Multi-Resol. PD	0.05
Total	0.9

❖ Power: 900µW

Measured Input-Jitter-Filtering Capability @2GHz

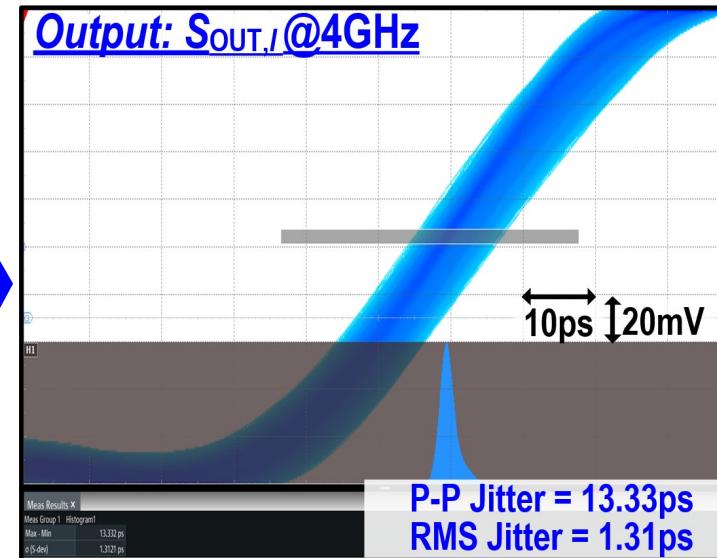
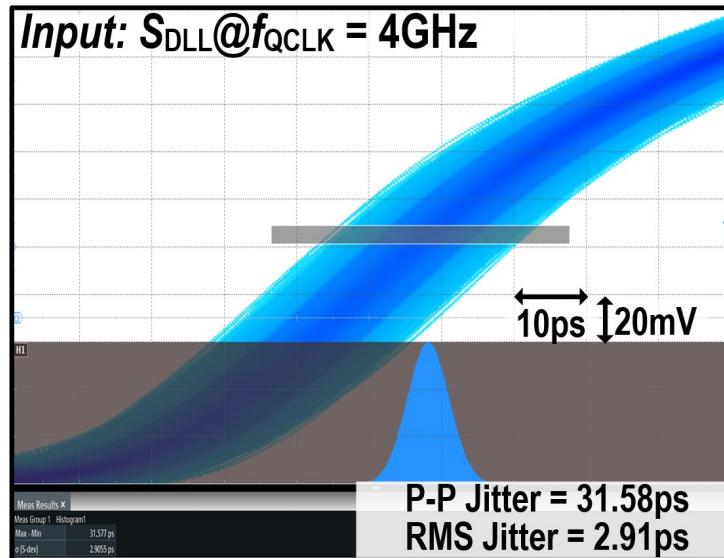


❖ Jitter_{RMS}: 2.94ps → 1.22ps

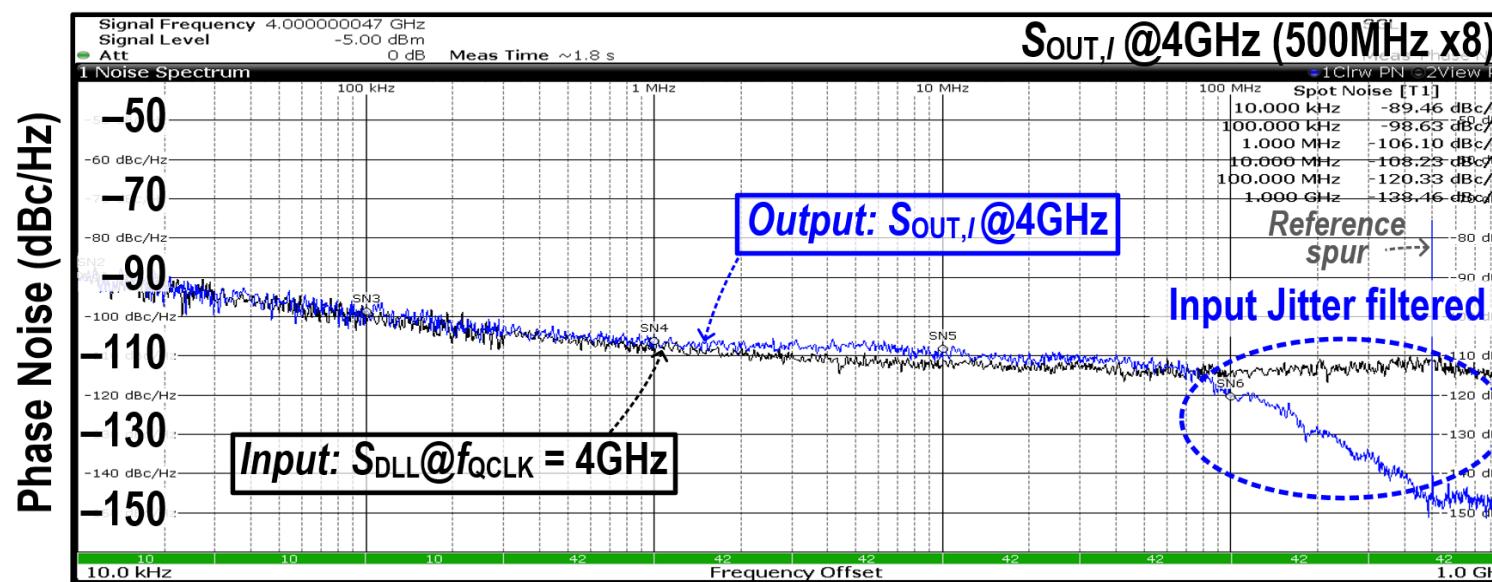


❖ Out-of-band PN of S_{DLL}
→ Low-pass filtered at $S_{\text{OUT},I}$

Measured Input-Jitter-Filtering Capability @4GHz

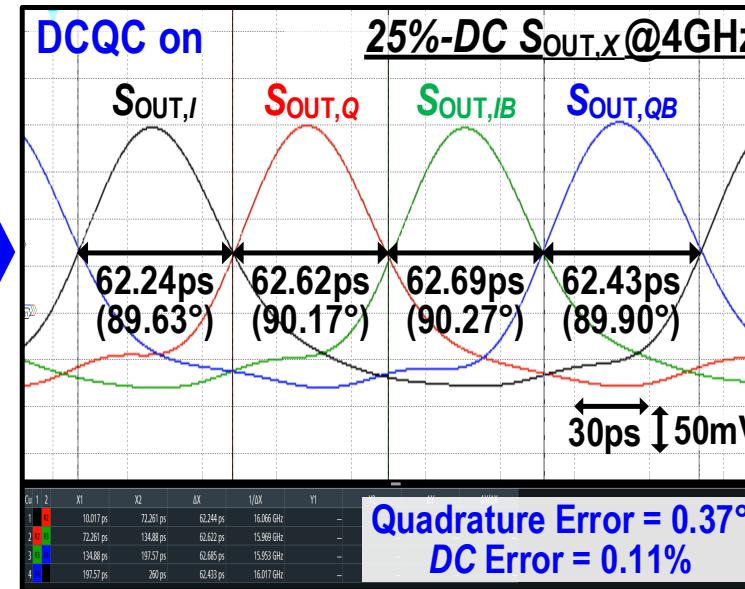
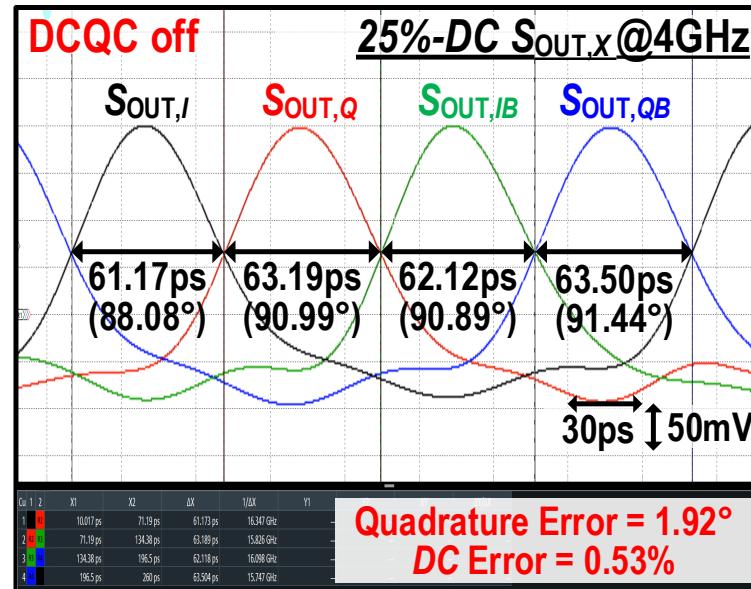
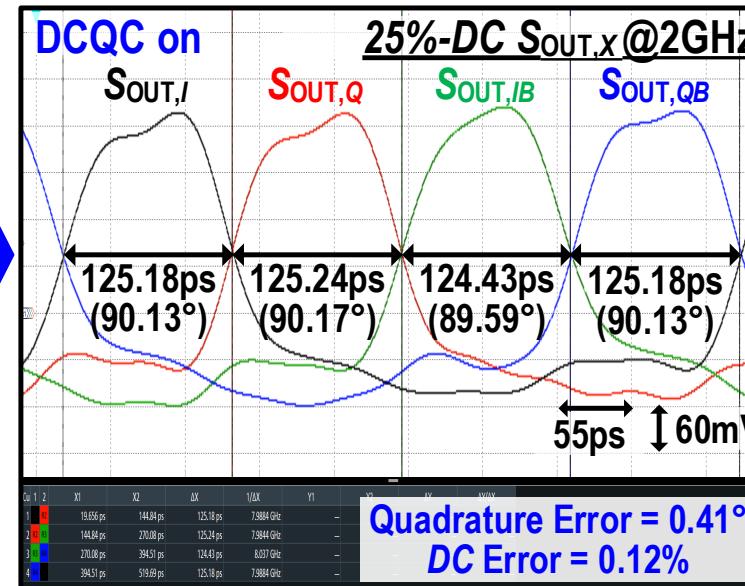
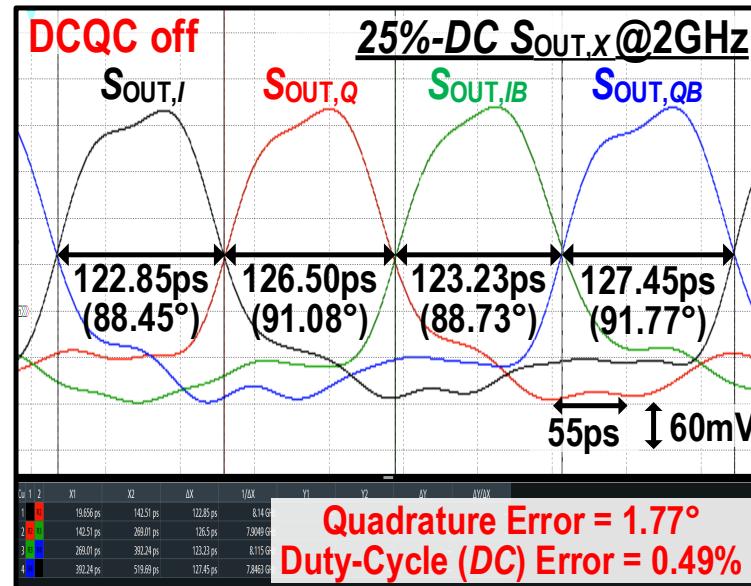


❖ Jitter_{RMS}: 2.91ps → 1.31ps



❖ Out-of-band PN of S_{DLL}
→ Low-pass filtered at $S_{\text{OUT},I}$

Measured Waveforms of Quad. Outputs @2 & 4GHz



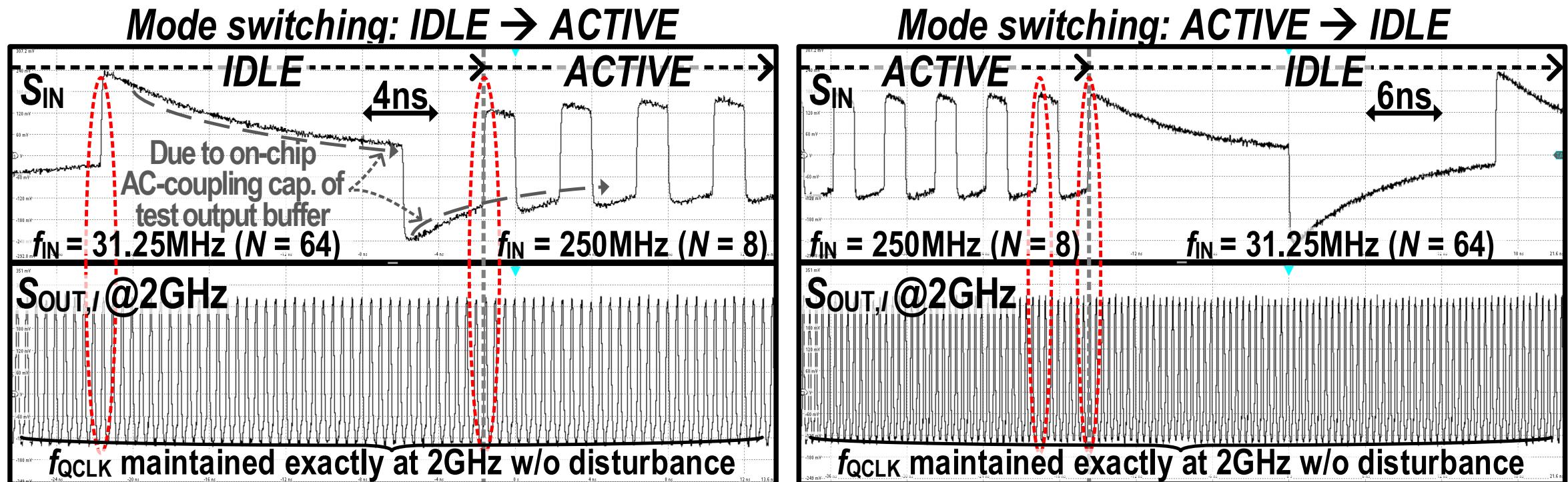
❖ Quadrature Error: **0.41°**

❖ Duty-Cycle Error: **0.12%**

❖ Quadrature Error: **0.37°**

❖ Duty-Cycle Error: **0.11%**

Measured Seamless Transition btw. *IDLE* and *ACTIVE* Modes



- ❖ $S_{OUT,I}$ maintains the same frequency **without experiencing any disturbances** despite abrupt transitions between *IDLE* and *ACTIVE* modes.

Performance Comparison

	This work	ISSCC'20 [1]	TCASII'17 [2]	TVLSI'19 [3]	ESSCIRC'21 [4]	JSSC'21 [5]
Process	40nm	40nm	65nm	55nm	28nm	28nm
Architecture	Digital-PLL QCG	Digital-DLL QEC	Digital-DLL QEC	Digital-DLL QEC	Digital-DLL QEC	Digital-DLL QCG
Quadrature Clock	Gen./ Correc.	Correc.	Correc.	Correc.	Correc.	Gen./ Correc.
Duty-Cycle (DC)	25% & 50%	50%	50%	50%	50%	50%
Freq. (f_{QCLK}) Range	1.0 – 4.0GHz	0.8 – 2.3GHz	1.25GHz	1.0 – 3.0GHz	0.8 – 3.2GHz	1.3 – 4.0GHz
Jitter Filtering	Yes	No	No	No	No	No
Input → Output Jitter _{rms} @ f_{QCLK}	2.94ps → 1.22ps @2.0GHz	2.28ps → 2.34ps @2.3GHz	1.84ps → 2.53ps @1.25GHz	1.85ps → 2.14ps @3.0GHz	NA* → 1.31ps @3.2GHz	0.96ps → 1.82ps @4.0GHz
Quadrature Error	< 0.5°	< 2.18°	< 0.48°	< 1.11°	< 1.84°	< 2.82°
Power Cons. @ f_{QCLK}	0.9mW@2.0GHz	8.9mW@2.3GHz	2.3mW@1.25GHz	2.1mW@3.0GHz	9.80mW@3.2GHz	6.5mW@4.0GHz
Power Efficiency	0.45mW/GHz	3.87mW/GHz	1.82mW/GHz	0.69mW/GHz	3.06mW/GHz	1.63mW/GHz
Active Area	0.011mm ²	0.012mm ² **	0.004mm ² **	0.003mm ²	0.010mm ²	0.004mm ²

* Input Jitter was not reported, ** Estimated from die micrograph

Conclusions

- ❖ An ultra-low-power clock distribution for high-speed DRAM interfaces w/ QCG
- ❖ Proposed low-power QCG:
 - Sub-mW Type-II DPLL architecture to filter the jitter of input clock
 - DCQC to generate 1-4GHz wide-frequency-range 25%-DC quadrature signals with a small quad. error of less than 0.5°
 - SMS-DIV to enable an immediate transition of modes

A 32-Gb/s/pin 0.51-pJ/b Single-Ended Resistor-less Impedance-Matched Transmitter with a T-Coil-Based Edge-Boosting Equalizer in 40nm CMOS

Jung-Hun Park¹, Hyeonseok Lee¹, Hoyeon Cho¹, Sanghee Lee¹,
Kwang-Hoon Lee¹, Han-Gon Ko², and Deog-Kyoong Jeong¹



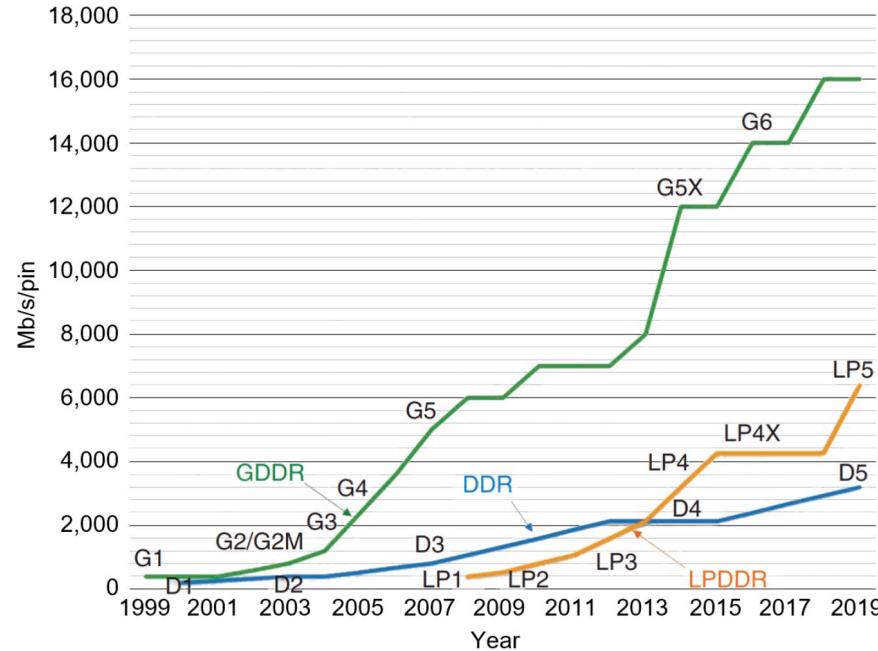
¹Seoul National University, Seoul, Korea

²ONEsemiconductor, Gyeonggi, Korea

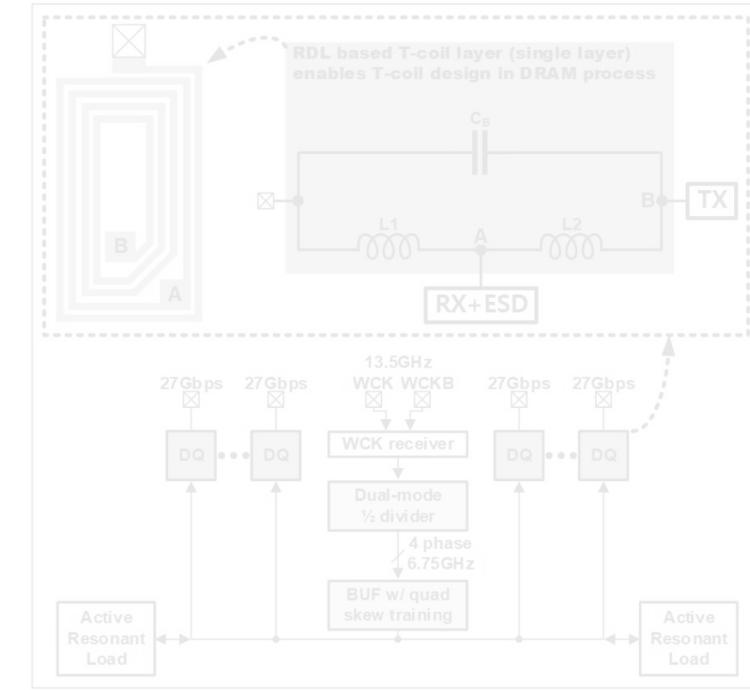
Outline

- **Introduction**
- **Prior Arts**
- **Proposed Driver and Equalizer**
 - PN-over-NP Driver
 - T-Coil-Based Edge-Boosting Equalizer
- **Circuit Implementation**
- **Measurement Results**
- **Conclusion**

Introduction



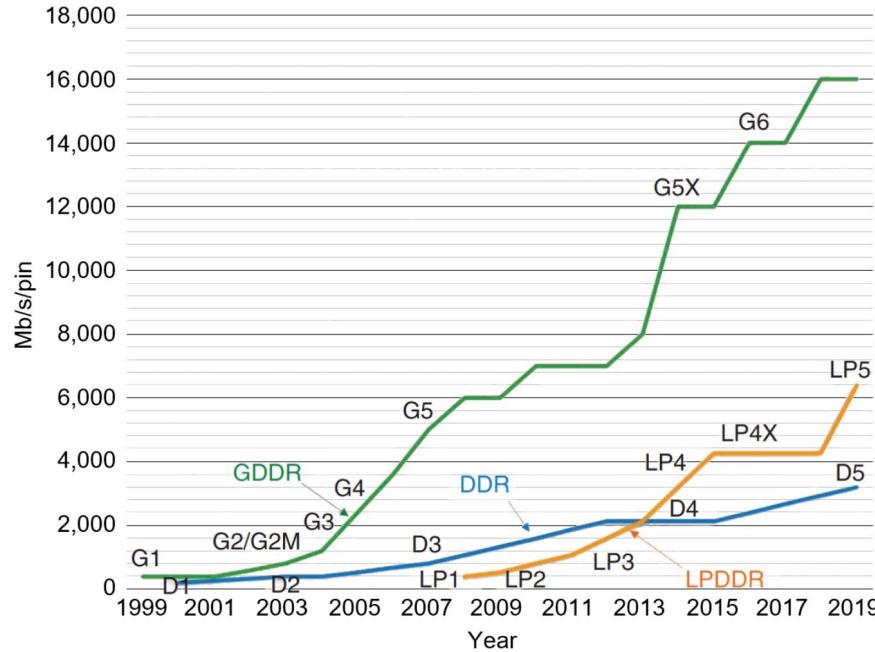
ISSC Magazine'19



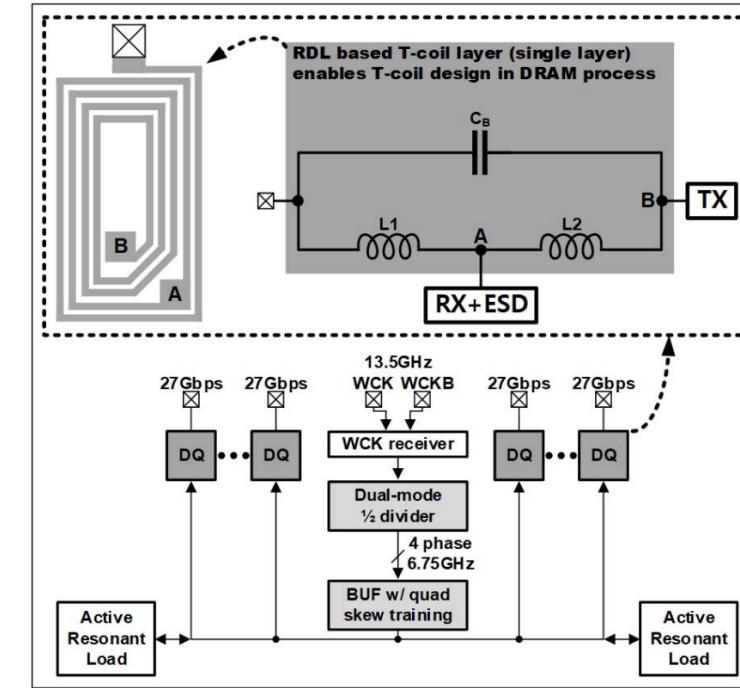
D. Lee, ISSCC'22 [1]

- **GDDR's per-pin bandwidth exceeds 20Gb/s/pin**
- **RDL-based T-coil is proposed to alleviate the effect of C_{load}**

Introduction



ISSC Magazine'19



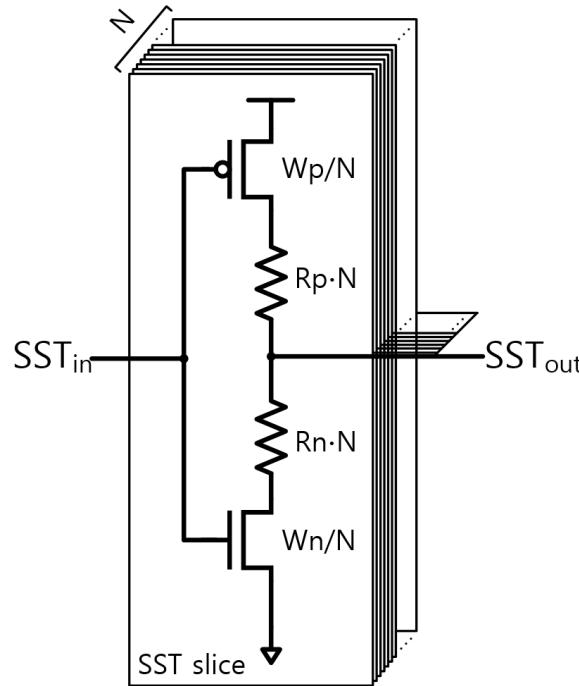
D. Lee, ISSCC'22 [1]

- GDDR's per-pin bandwidth exceeds 20Gb/s/pin
- RDL-based T-coil is proposed to alleviate the effect of C_{load}

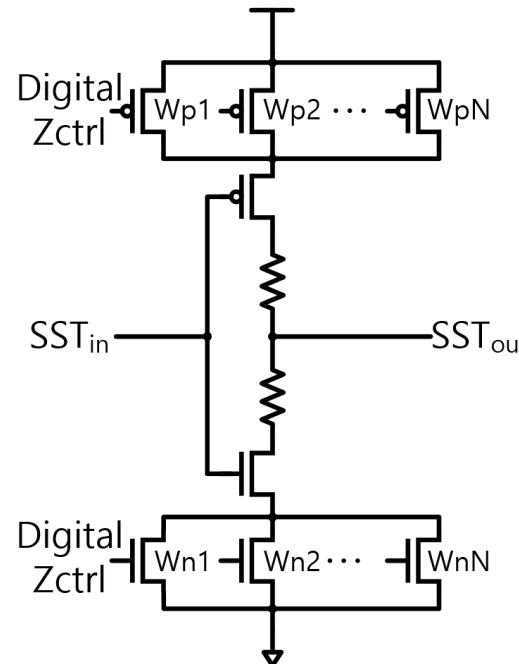
Outline

- **Introduction**
- **Prior Arts**
- **Proposed Driver and Equalizer**
 - PN-over-NP Driver
 - T-Coil-Based Edge-Boosting Equalizer
- **Circuit Implementation**
- **Measurement Results**
- **Conclusion**

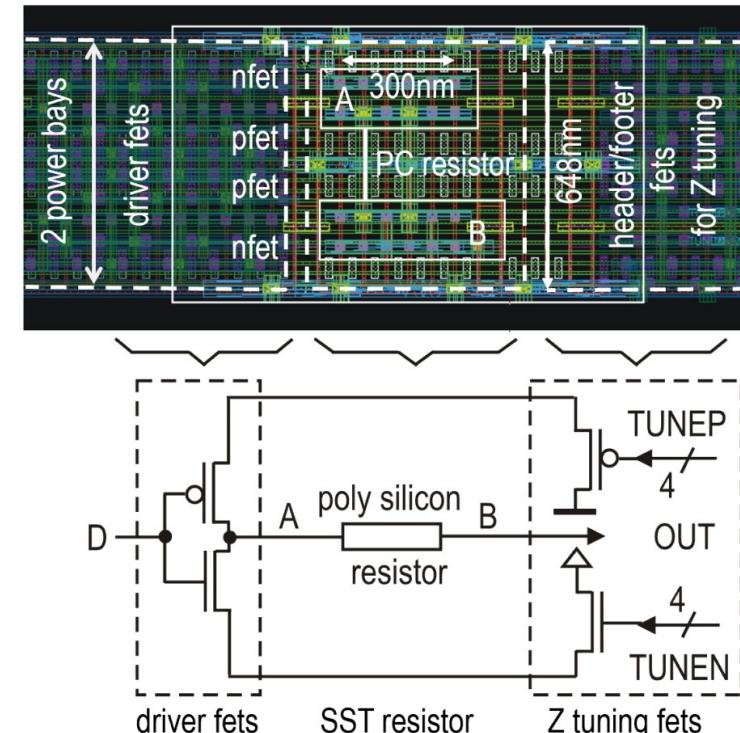
Prior Arts – Drivers



SST driver



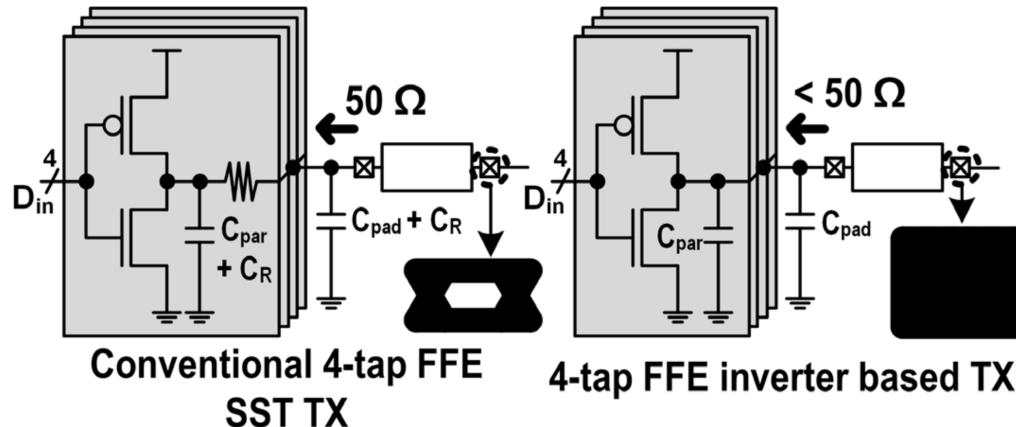
M. Kossel, JSSC'08



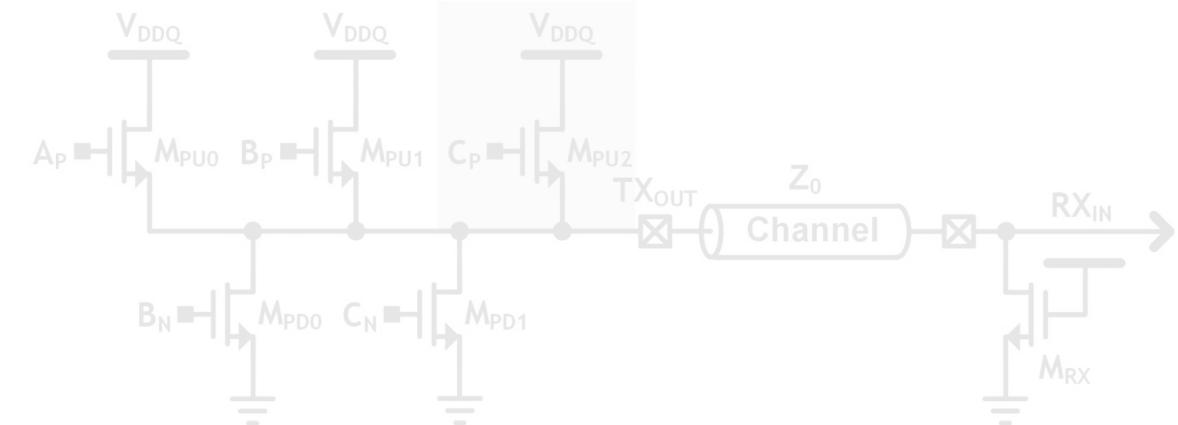
M. Kossel, ISSCC'21

- ☺ Great linearity and impedance matching
- ☺ Large output swing
- ☹ Large area due to series resistors
- ☹ TR size $\uparrow \rightarrow$ power consumption \uparrow

Prior Arts – Drivers



C. Moon, ISSCC'22 [2]



Y.-U. Jeong, JSSC'21 [3]

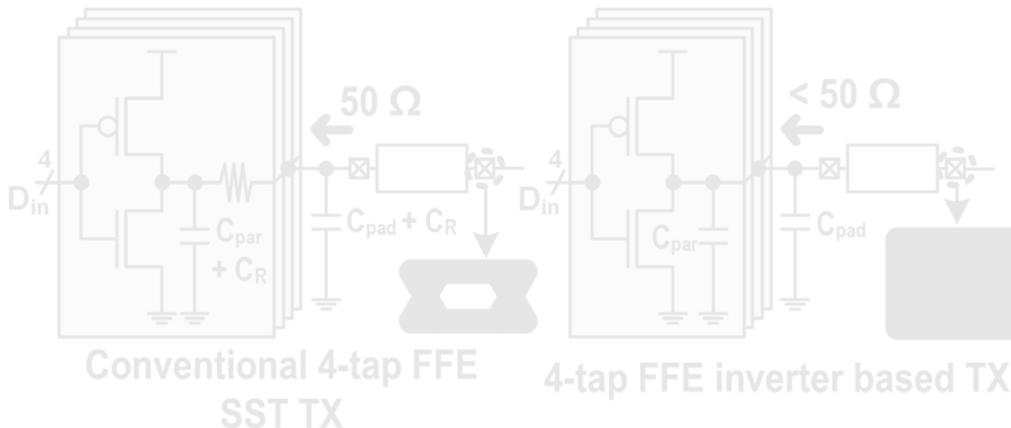
■ Inverter without series resistor

- ☺ Very small area
- ☹ No TX-matching, only far-end matching

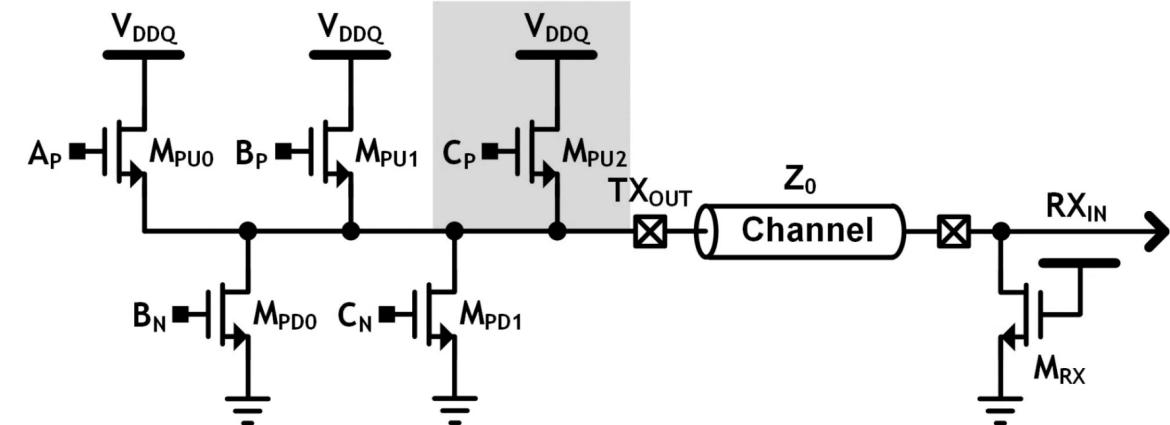
■ Impedance-matched PAM-4 driver

- ☺ Output impedance is matched at four PAM-4 levels
- ☹ Require additional encoder and complicate ZQ calibration loops

Prior Arts – Drivers



C. Moon, ISSCC'22 [2]



Y.-U. Jeong, JSSC'21 [3]

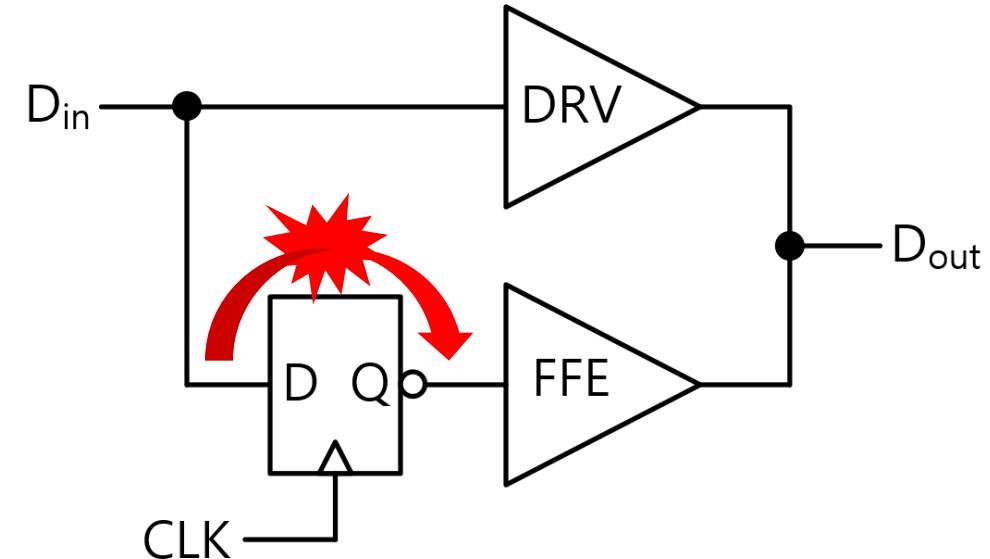
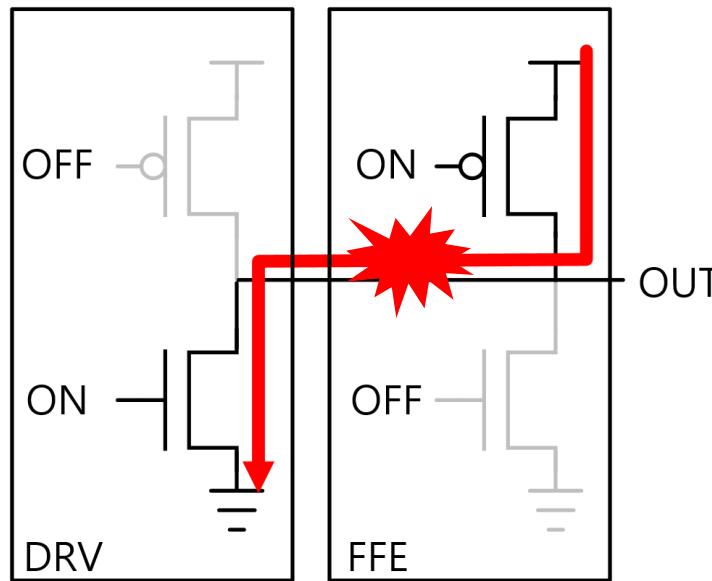
Inverter without series resistor

- 😊 Very small area
- 😢 No TX-matching, only far-end matching

Impedance-matched PAM-4 driver

- 😊 Output impedance is matched at four PAM-4 levels
- 😢 Require additional encoder and complicate ZQ calibration loops

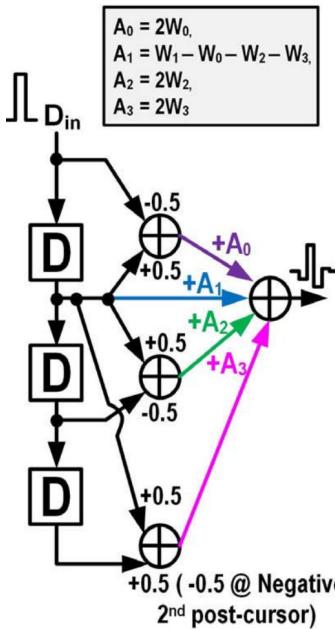
Prior Arts – Equalizers



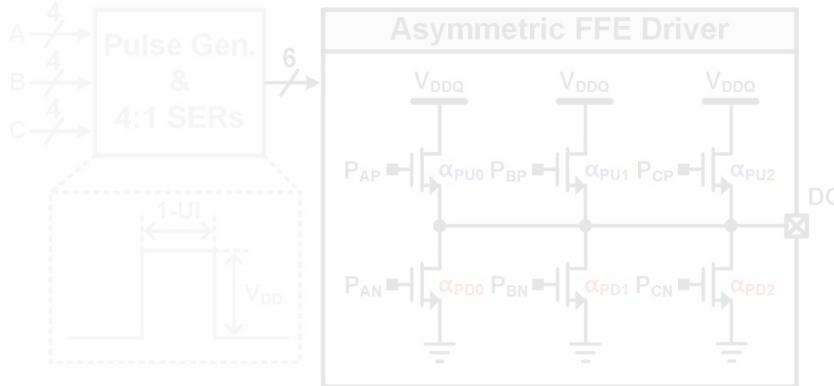
■ De-emphasis FFE

- 😊 Robust equalization
- 😊 Good impedance matching
- 😢 Wasted static current when there is no data transition
- 😢 1-UI delayed data is required

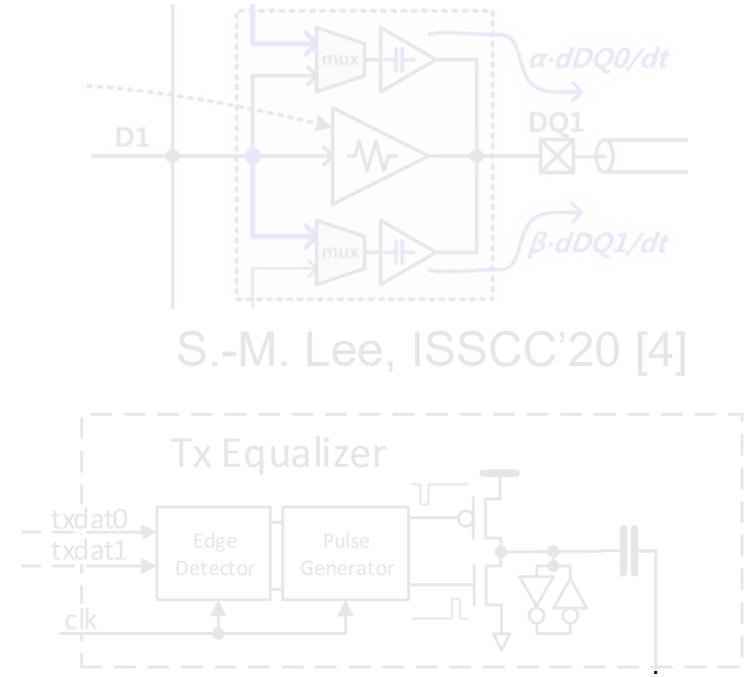
Prior Arts – Equalizers



C. Moon,
ISSCC'22 [2]



Y.-U. Jeong, JSSC'21 [3]



J. M. Wilson, ISSCC'18 [5]

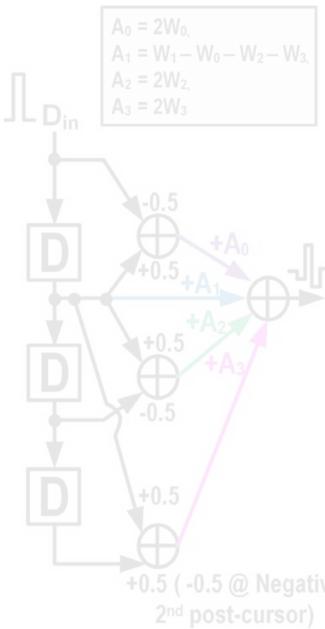
■ Addition-only FFE

- ☺ Only additions between taps → better power efficiency
- ☹ Sub-filters composed of logic gates consume power

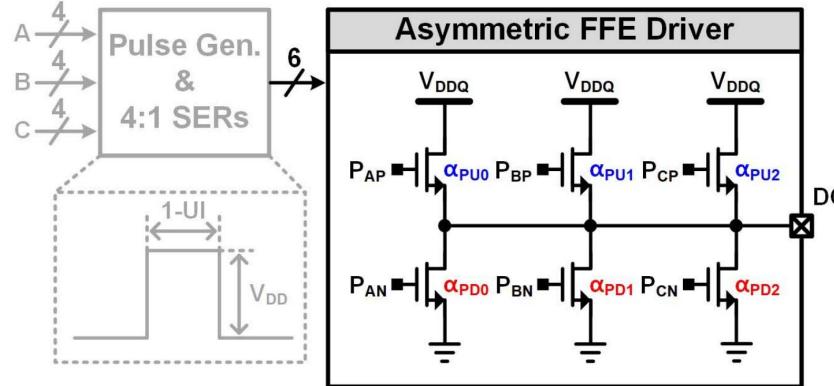
■ Pre-emphasis equalizer with pulse generator

- ☺ No wasted current → better power efficiency
- ☹ Bad impedance matching during transition → additional HW (pulse gen...)

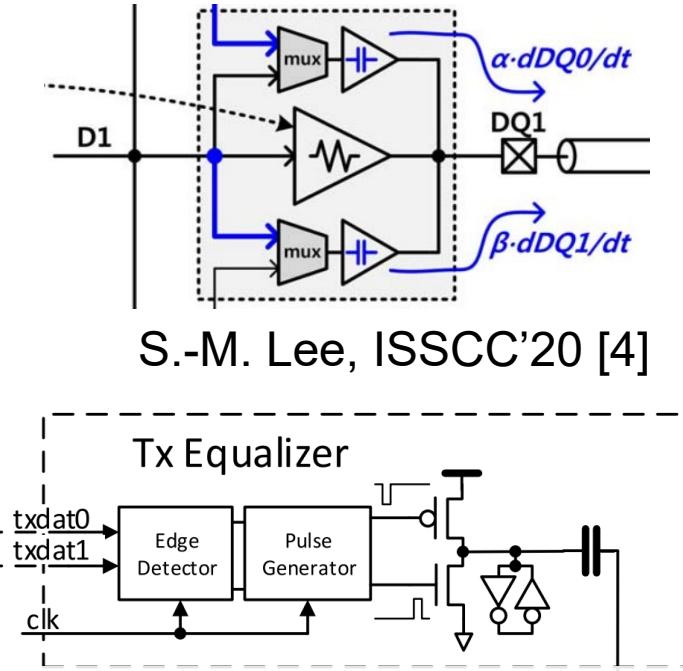
Prior Arts – Equalizers



C. Moon,
ISSCC'22 [2]



Y.-U. Jeong, JSSC'21 [3]



J. M. Wilson, ISSCC'18 [5]

Addition-only FFE

- ☺ Only additions between taps → better power efficiency
- ☹ Sub-filters composed of logic gates consume power

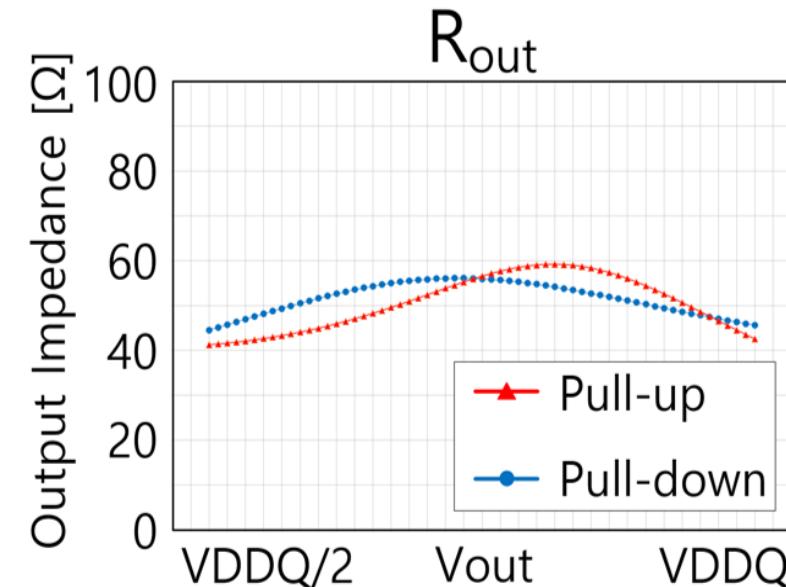
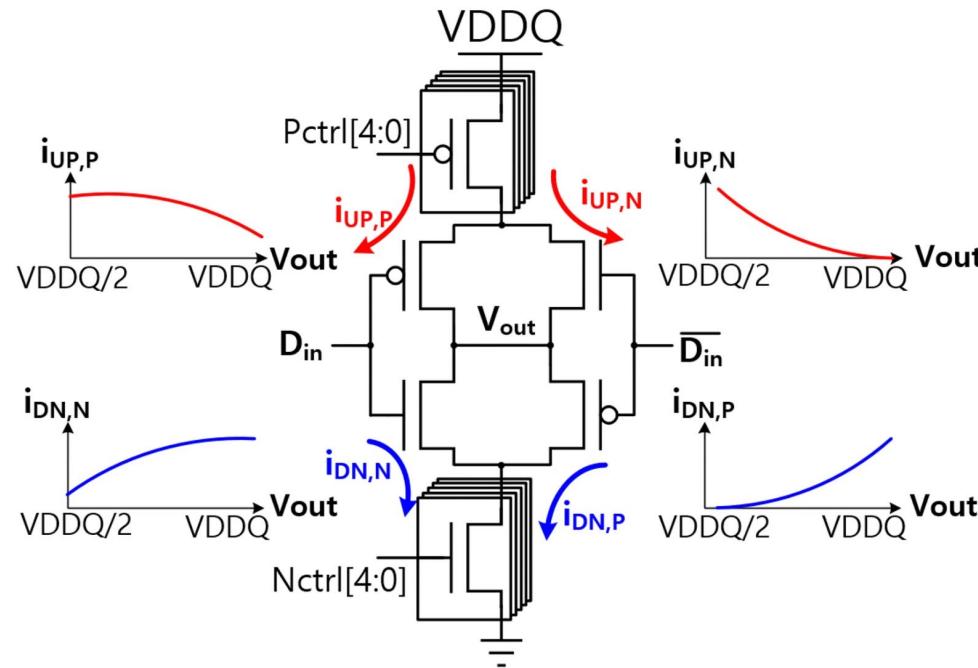
Pre-emphasis equalizer with pulse generator

- ☺ No wasted current → better power efficiency
- ☹ Bad impedance matching during transition → additional HW (pulse gen...)

Outline

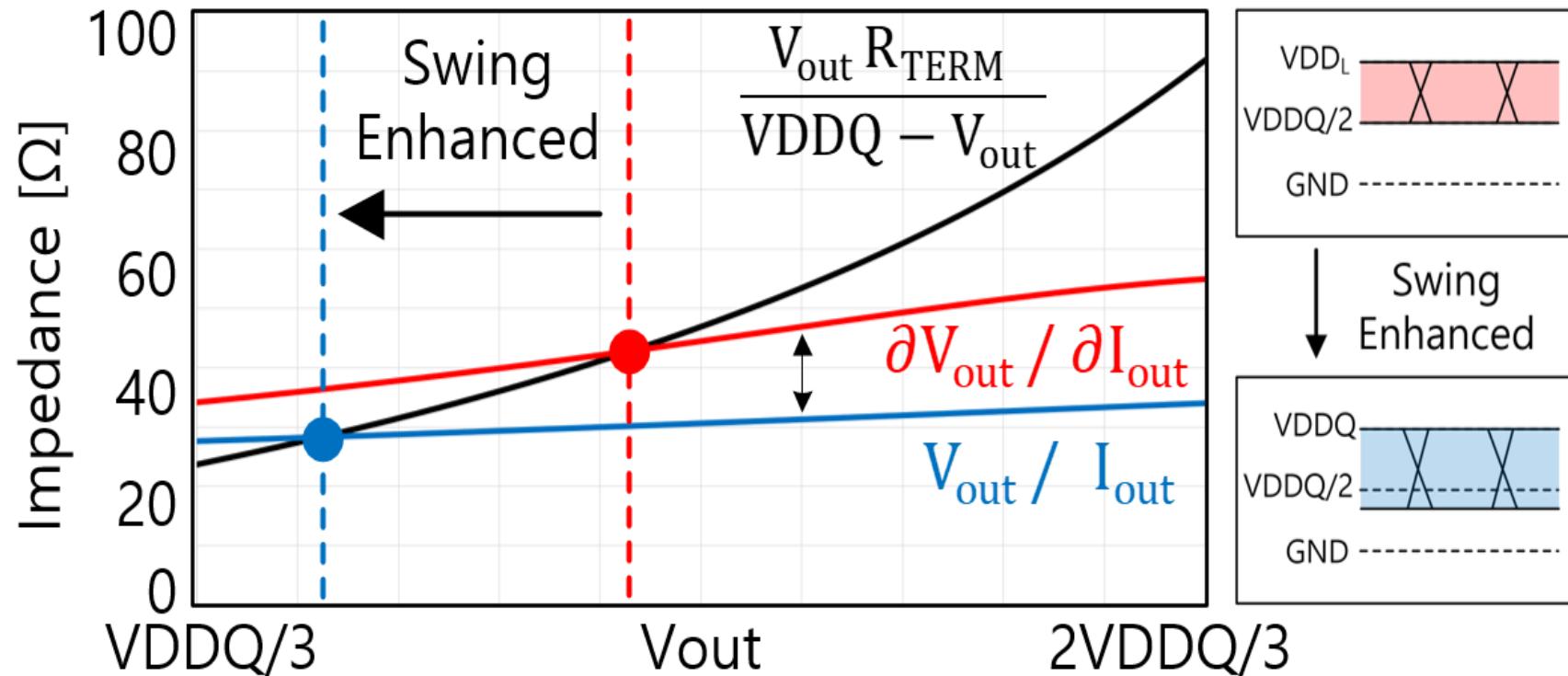
- Introduction
- Prior Arts
- **Proposed Driver and Equalizer**
 - PN-over-NP Driver
 - T-Coil-Based Edge-Boosting Equalizer
- Circuit Implementation
- Measurement Results
- Conclusion

PN-over-NP Driver



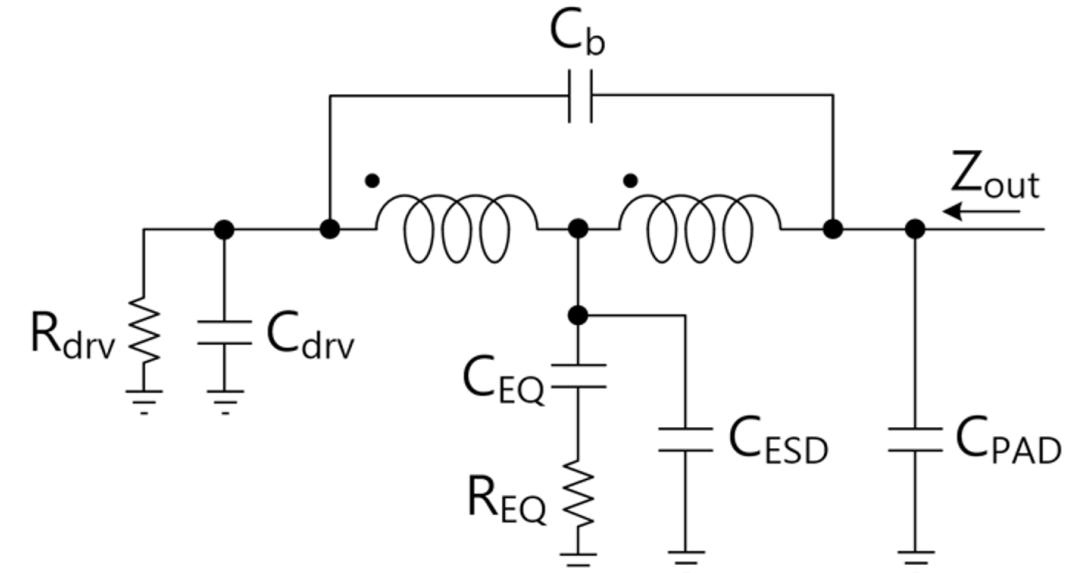
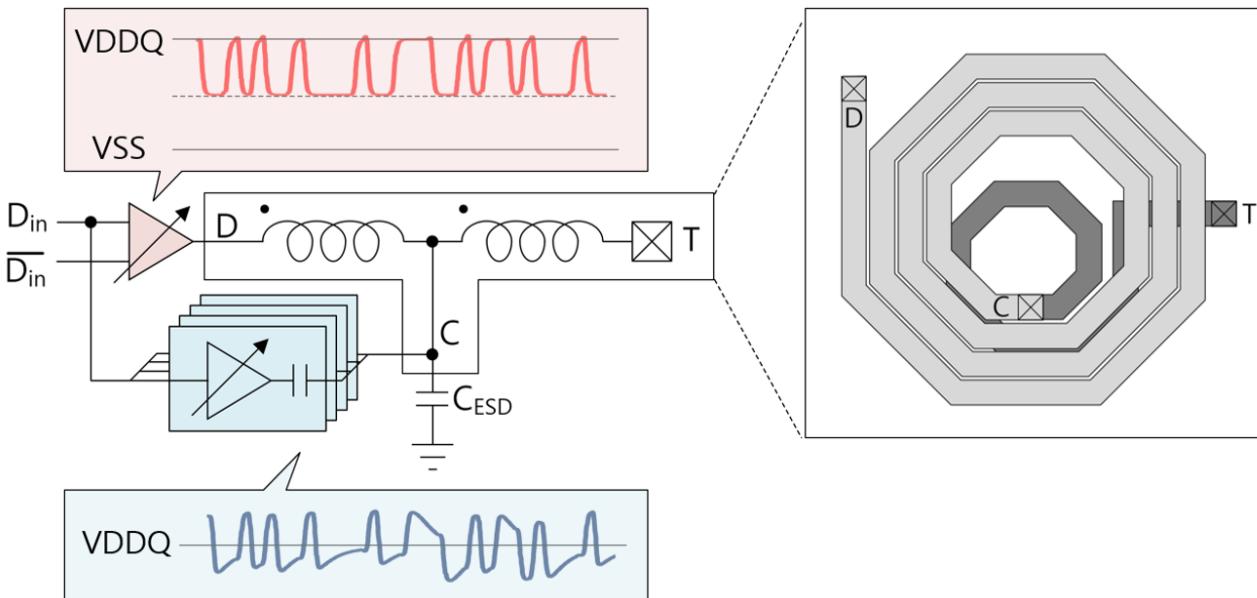
- Complementary N-over-P driver
- Output impedance is matched to within $50 \pm 10 \Omega$
- 5-b thermometer control for PVT tuning

PN-over-NP Driver



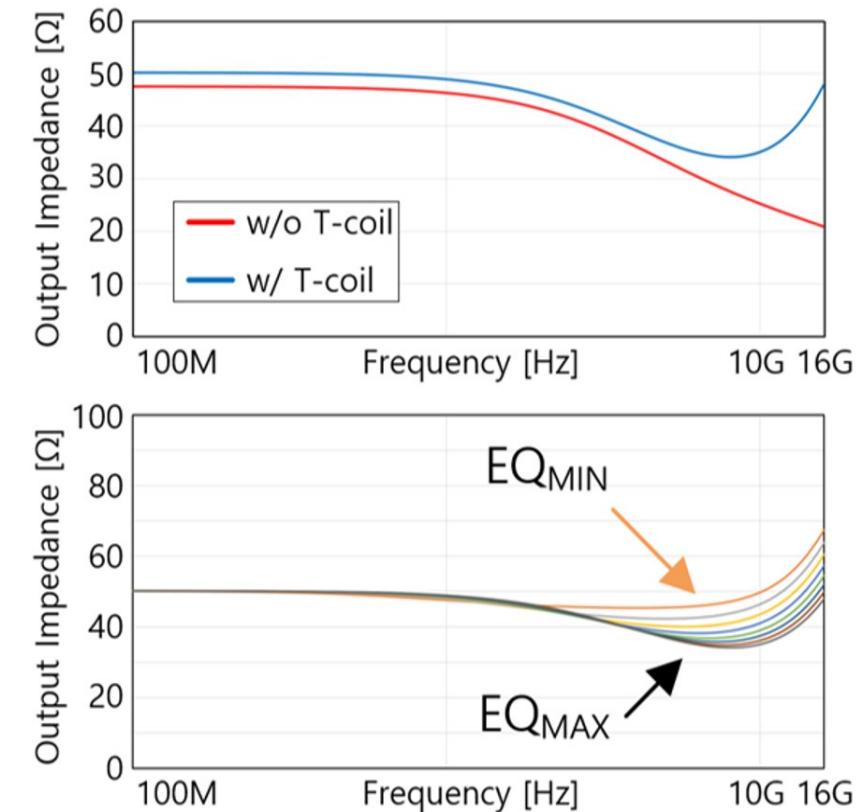
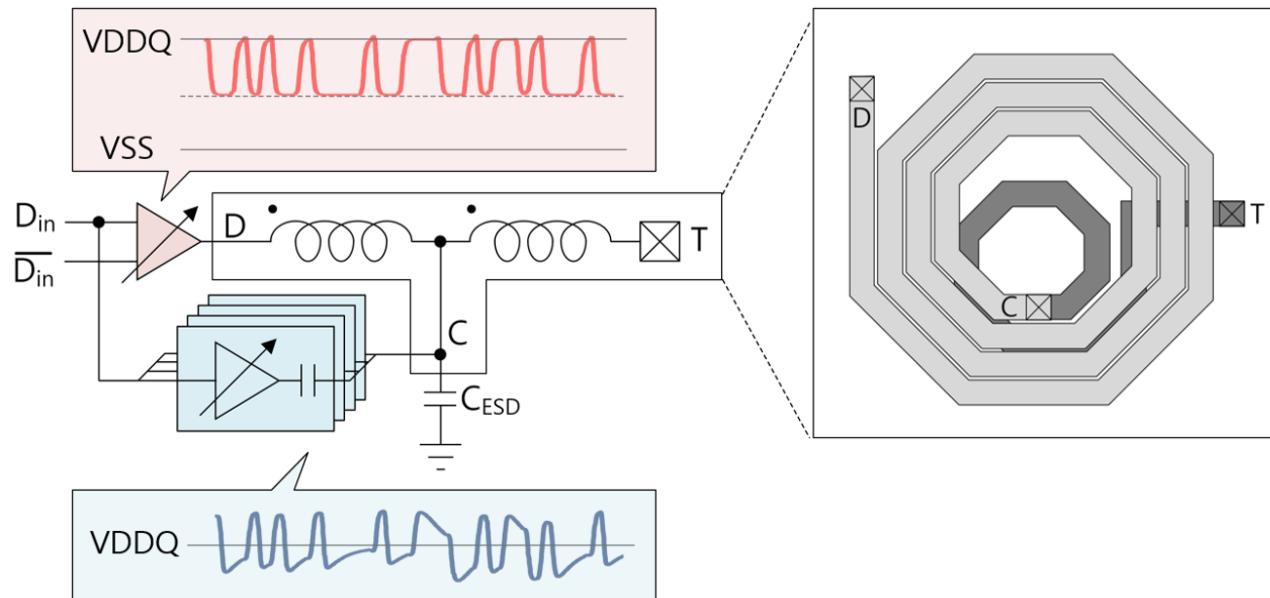
- Difference between small-signal Z_{out} and large-signal Z_{out}
 - Partially linear impedance characteristics → small large-signal impedance
- Output swing is enhanced due to the small V_{out}/I_{out} at LOW

T-coil-Based Edge-Boosting Equalizer



- T-coil is designed with top metal layer emulating RDL
- Edge-boosting EQ is connected to the center tap of the T-coil

T-coil-Based Edge-Boosting Equalizer

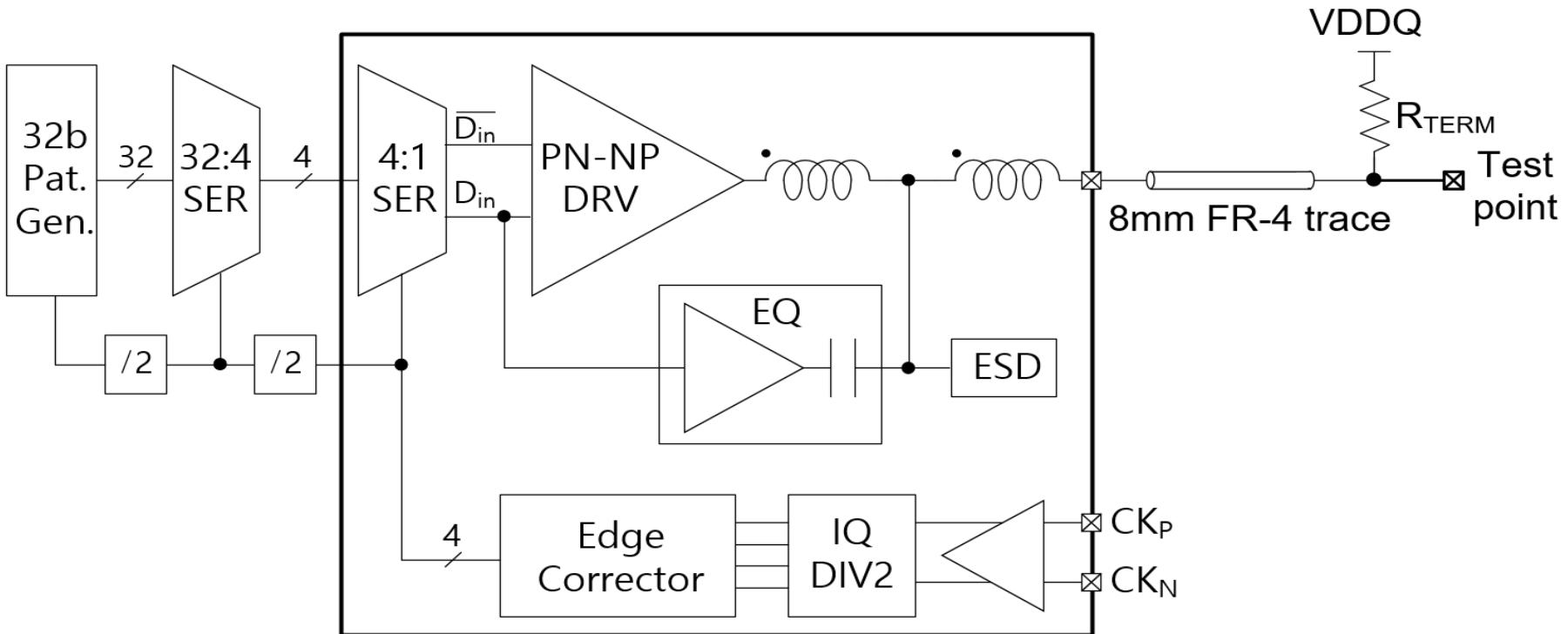


- The impedance drop rate is improved by 47% at high freq.
- EQ strength does not significantly affect output impedance

Outline

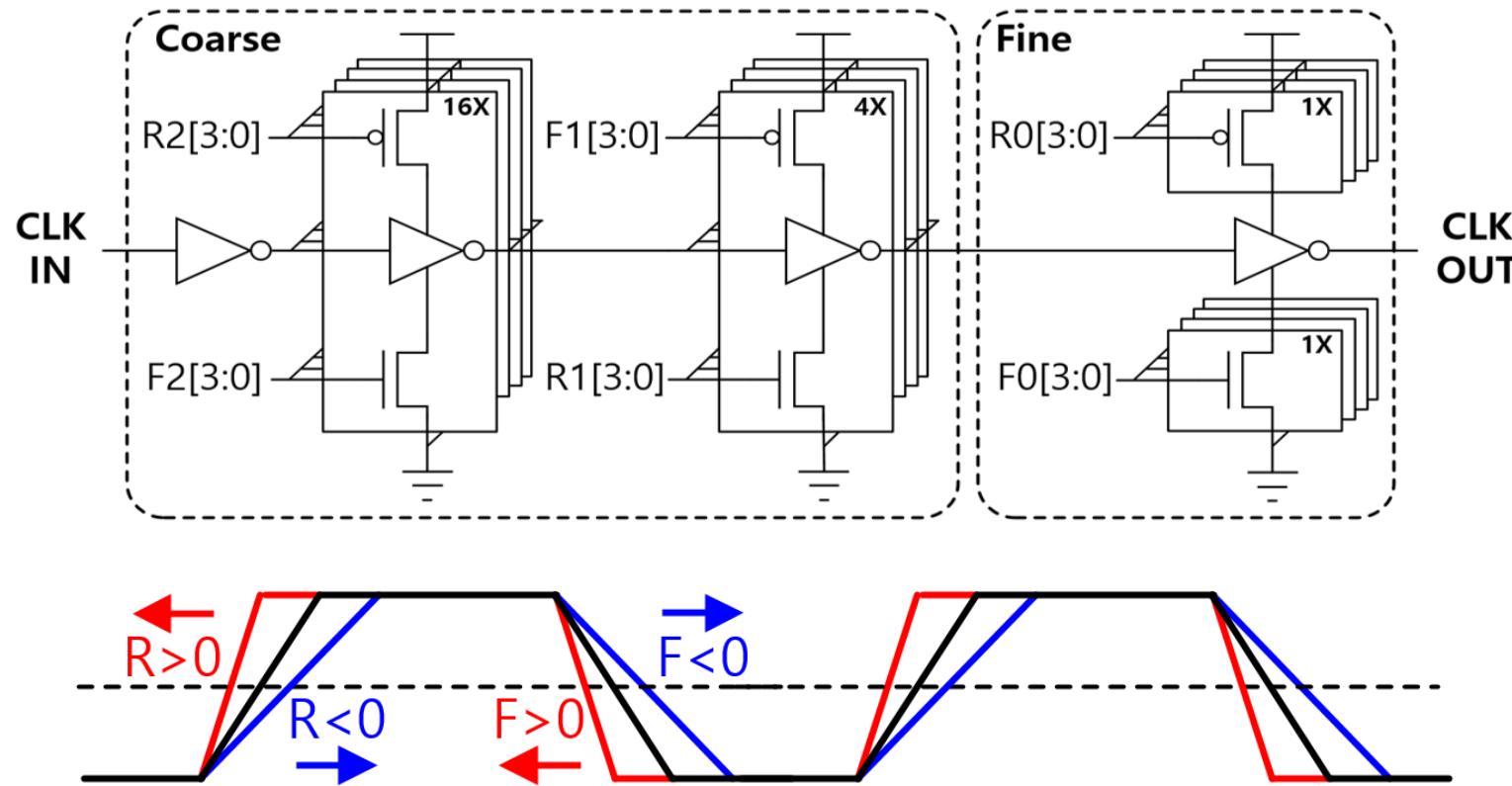
- Introduction
- Prior Arts
- Proposed Driver and Equalizer
 - PN-over-NP Driver
 - T-Coil-Based Edge-Boosting Equalizer
- Circuit Implementation
- Measurement Results
- Conclusion

Proposed Transmitter



- Includes 4:1 serializer, driver, equalizer, and clock path
- 16-GHz external differential clock
- VDDQ termination

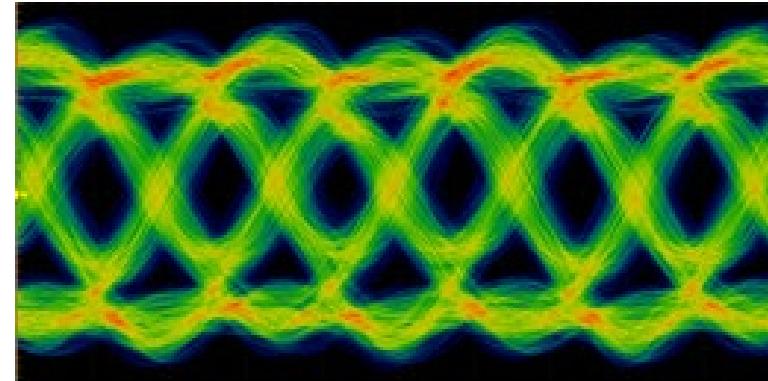
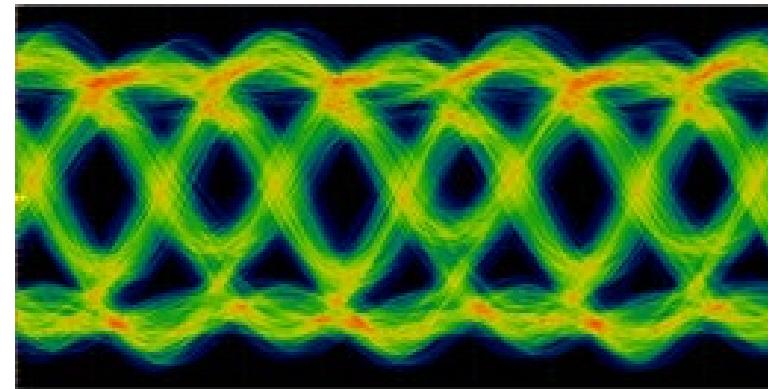
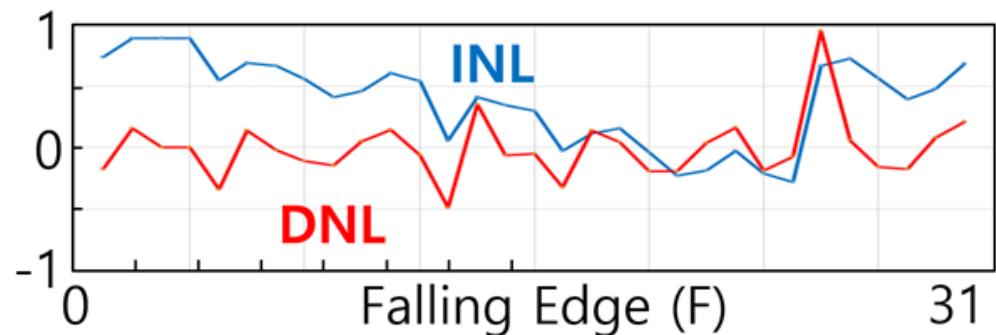
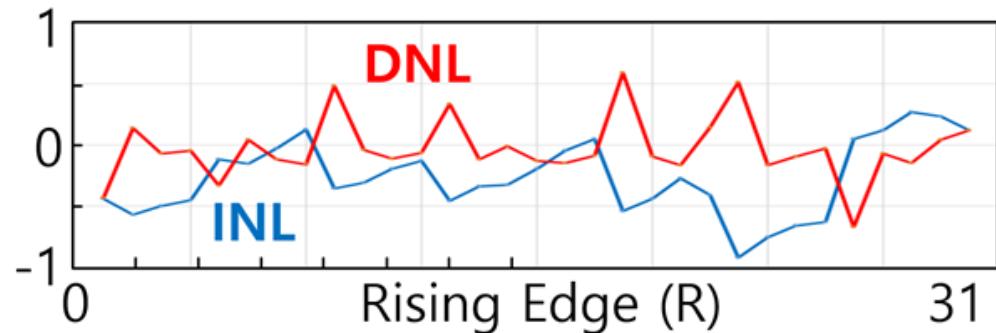
Clock Error Corrector



Control Code	Duty	Phase
$R > F > 0$	Up	Lead
$R = 0 > F$	Up	-
$0 > R > F$	Up	Lag
$R = F > 0$	-	Lead
$R = F = 0$	-	-
$0 > R = F$	-	Lag
$0 < R < F$	Down	Lead
$R < F = 0$	Down	-
$0 > R > F$	Down	Lag

- CMOS-based clock error corrector
- Duty-cycle and phase are adjusted simultaneously

Clock Error Corrector

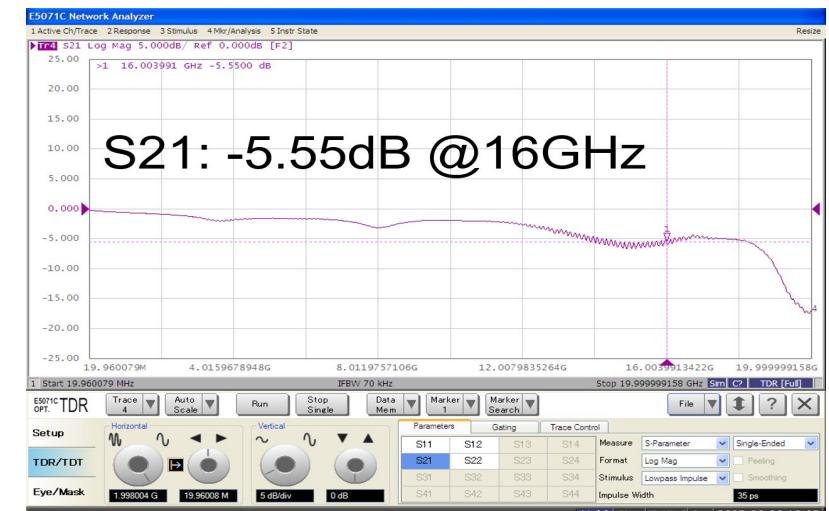
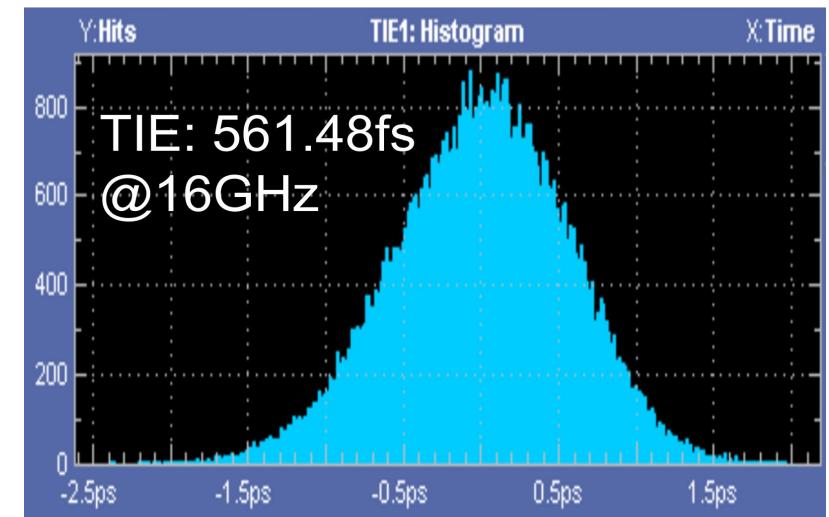
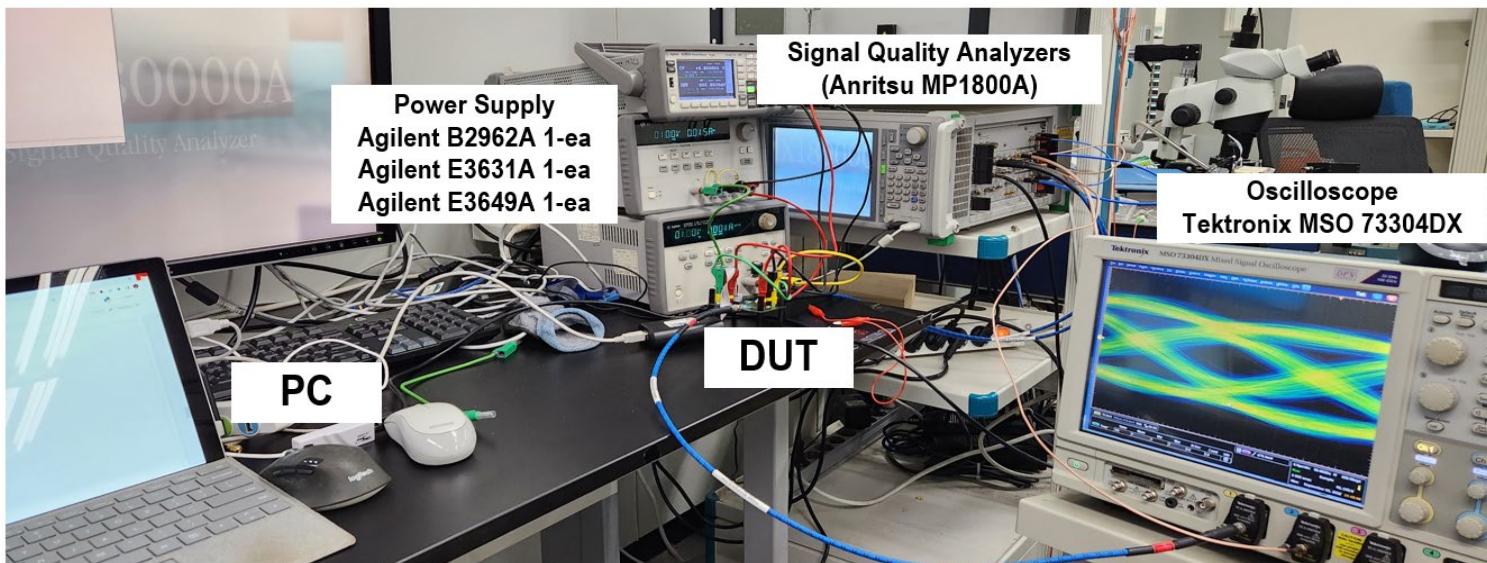
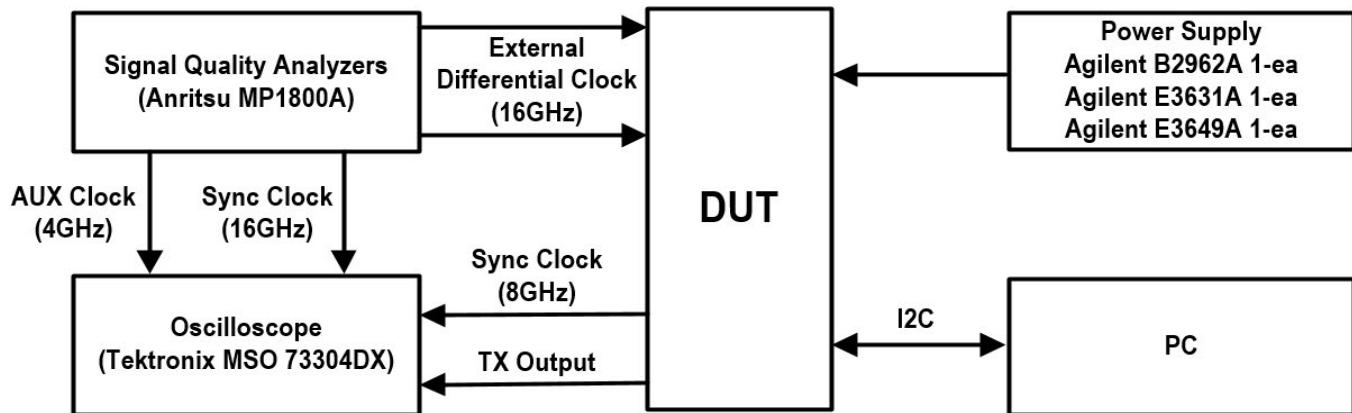


- INL & DNL < ± 1 LSB (post-layout simulation)
- Effectiveness is verified by actual measurements

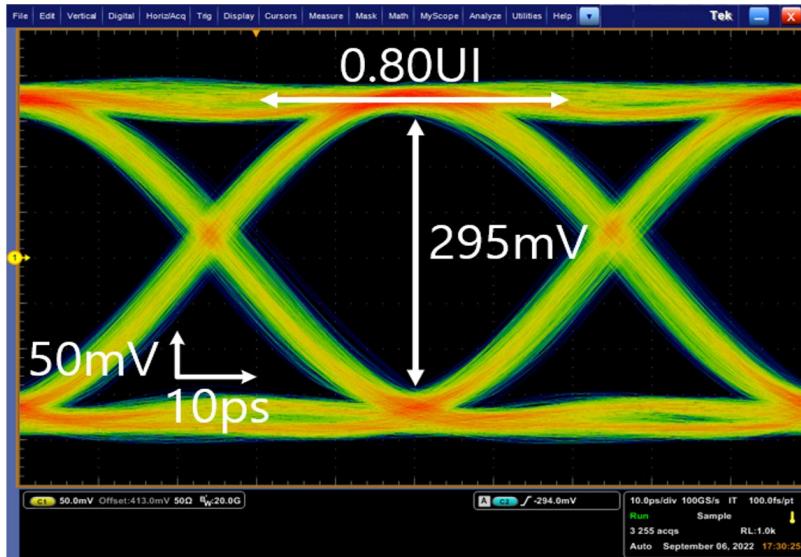
Outline

- **Introduction**
- **Prior Arts**
- **Proposed Driver and Equalizer**
 - PN-over-NP Driver
 - T-Coil-Based Edge-Boosting Equalizer
- **Circuit Implementation**
- **Measurement Results**
- **Conclusion**

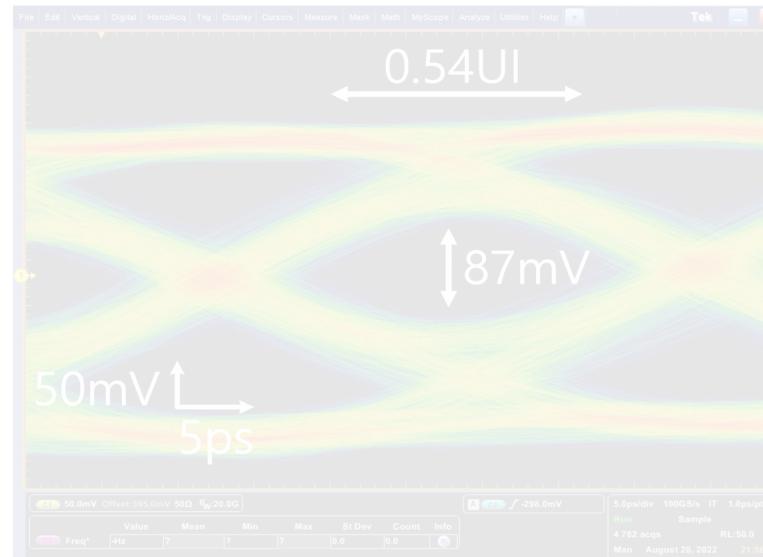
Measurement Setup



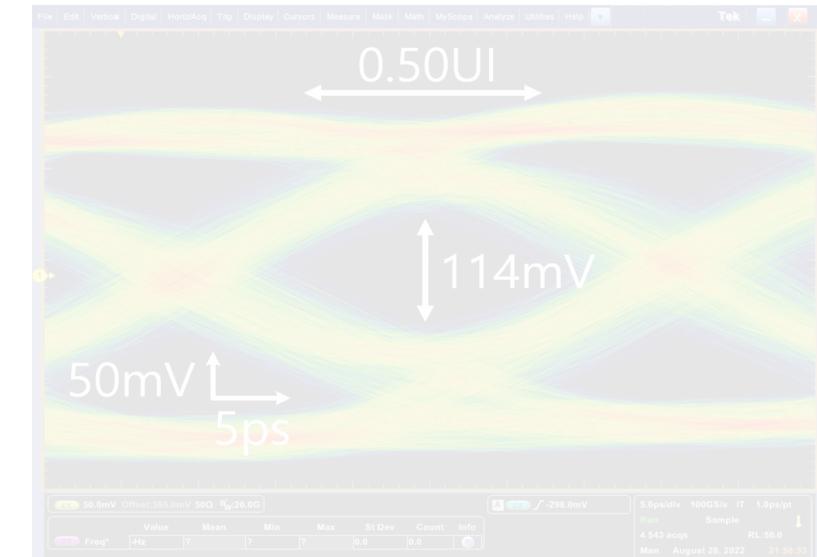
Eye Diagrams



w/o EQ @20.0Gb/s



w/o EQ @32.0Gb/s



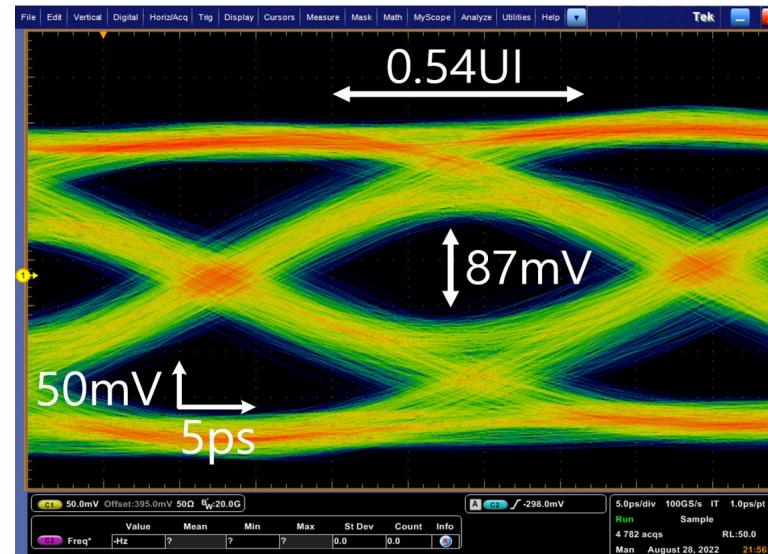
w/ EQ @32.0Gb/s

- Increased output swing enables large vertical eye margin
- Equalizer enables vertical eye margin > 100mV at 32Gb/s

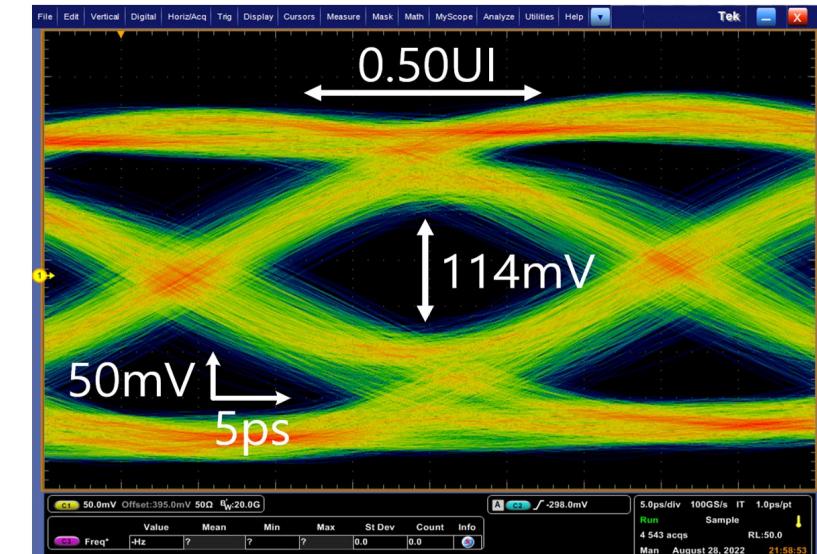
Eye Diagrams



w/o EQ @20.0Gb/s



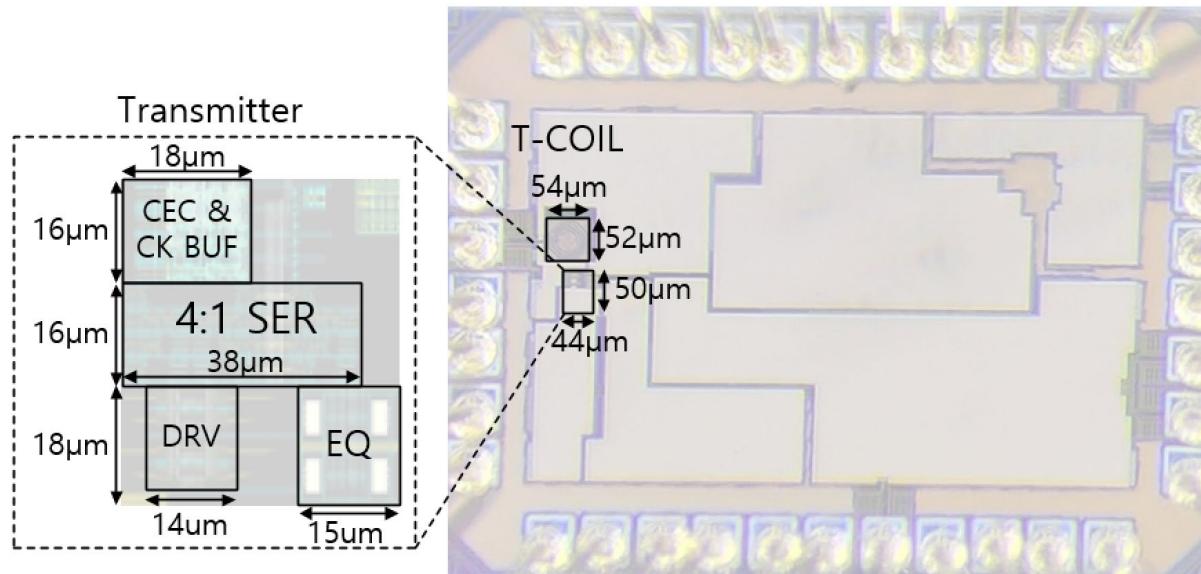
w/o EQ @32.0Gb/s



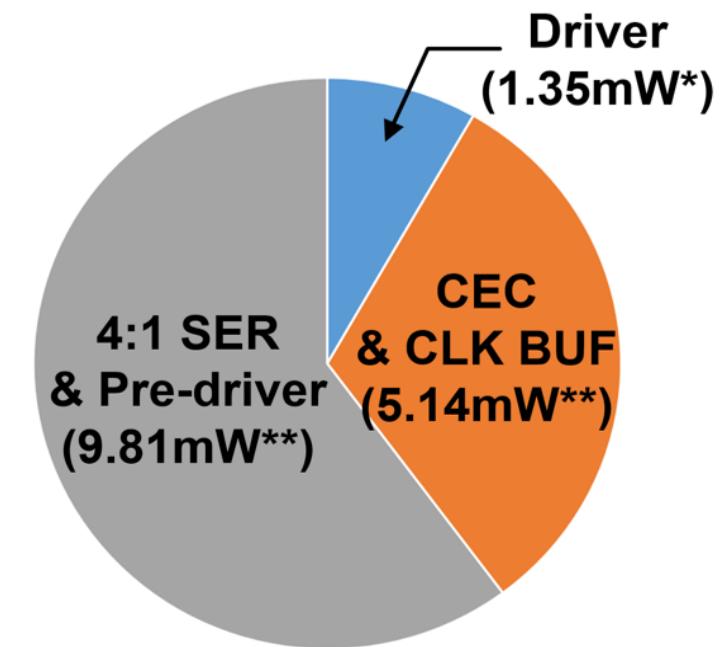
w/ EQ @32.0Gb/s

- Increased output swing enables large vertical eye margin
- Equalizer enables vertical eye margin > 100mV at 32Gb/s

Area and Power Breakdown



T-coil		54μm X 52μm
Transmitter	CEC & CK BUF	18μm X 16μm
	4:1 SER	38μm X 16μm
	DRV	14μm X 18μm
	EQ	15μm X 18μm



* Includes the VDDQ termination

** Measurement result is separated based on post-layout simulation results

- **5008um² including T-coil**
- **0.51pJ/b @32Gb/s**

Comparison Table

		ISSCC'22 [2]	JSSC'21 [3]	ISSCC'20 [4]	ISSCC'18 [5]	Kang JSSC'22	Ko ISSCC'20	Chiu ISSCC'20	This Work
Technology		28nm LPP	65nm CMOS	8nm FinFET	16nm FinFET	28nm CMOS	65nm CMOS	65nm CMOS	40nm CMOS
Data rate [Gb/s]		20	28	18	25	21	4	32	32
Signaling		NRZ	PAM-4	NRZ	GRS	Duobinary	NRZ	PAM-4	NRZ
Supply voltage	VDD [V]	1.1	1.0	0.85	0.75	1.0	1.2	1.2	1.0
	VDDQ [V]		0.6	1.35		0.8			0.6
Driver & Equalizer	Driver type	Inverter	N-over-N	High-voltage SST	Charge Pump	SST	Inverter	SST	PN-over-NP
	TX equalization	4-tap AFFE (pre & post 2)	2-tap pre-emphasis	2-tap Edge boosting	2-tap Edge boosting	3-tap FFE (pre & post)	2-tap FFE (post + XT)	3-tap FFE (half-rate)	2-tap Edge boosting
	No static current during IDLE state	O	O	O	X	X	X	X	O
	Impedance matching during transition	X	X	X	X	O	X	O	O
Energy efficiency [pJ/b]		1.18	0.58*	N/A	1.17**	0.67	0.9	0.97**	0.51
Area [mm ²]		0.00115	0.033	4.15	0.0102**	0.0072	0.0027***	0.009**	0.00501

* Excludes PRBS generator and 32:8 serializer (according to the power breakdown)

** TRX

*** Area / (# of I/O)

Outline

- Introduction
- Prior Arts
- Proposed Driver and Equalizer
 - PN-over-NP Driver
 - T-Coil-Based Edge-Boosting Equalizer
- Circuit Implementation
- Measurement Results
- Conclusion

Conclusion

- A 32-Gb/s single-ended transmitter has been proposed
 - PN-over-NP driver
 - Resistor-less design enables small area and better power efficiency
 - Large output swing due to the small large-signal impedance
 - T-coil-based edge-boosting equalizer
 - No current waste during idle data period
 - Output impedance is maintained even at high frequencies
- TX consumes the best power efficiency (0.51pJ/b) compared to state-of-the-arts

Thank You

A 1.1-V 6.4-Gb/s/pin 24-Gb DDR5 SDRAM with a Highly-Accurate Duty Corrector and NBTI-Tolerant DLL

Daehyun Kwon, Heon Su Jeong, Jaemin Choi, Wijong Kim, Jae Woong Kim,
Junsuh Yoon, Jungmin Choi, Sanguk Lee, Hyunsub Norbert Rie, Jin-il Lee,
Jongbum Lee, Taeseong Jang, JunHyung Kim, Sanghee Kang, Jungbum Shin,
Yanggyoon Loh, Chang Yong Lee, Junmyung Woo, Hyeseung Yu,
Changhyun Bae, Reum Oh, Young-soo Sohn, Changsik Yoo, Jooyoung Lee



Outline

- **Introduction of DDR5**
- **Key schemes**
 - GIO schemes for Low power consumption
 - NBTI tolerable schemes for DLL
 - Adaptive Body Bias for process variations
 - Highly-accurate duty cycle detector
 - High-speed schemes for transmitter
- **Implementation and measurements**
- **Conclusion**

Outline

■ Introduction of DDR5

■ Key schemes

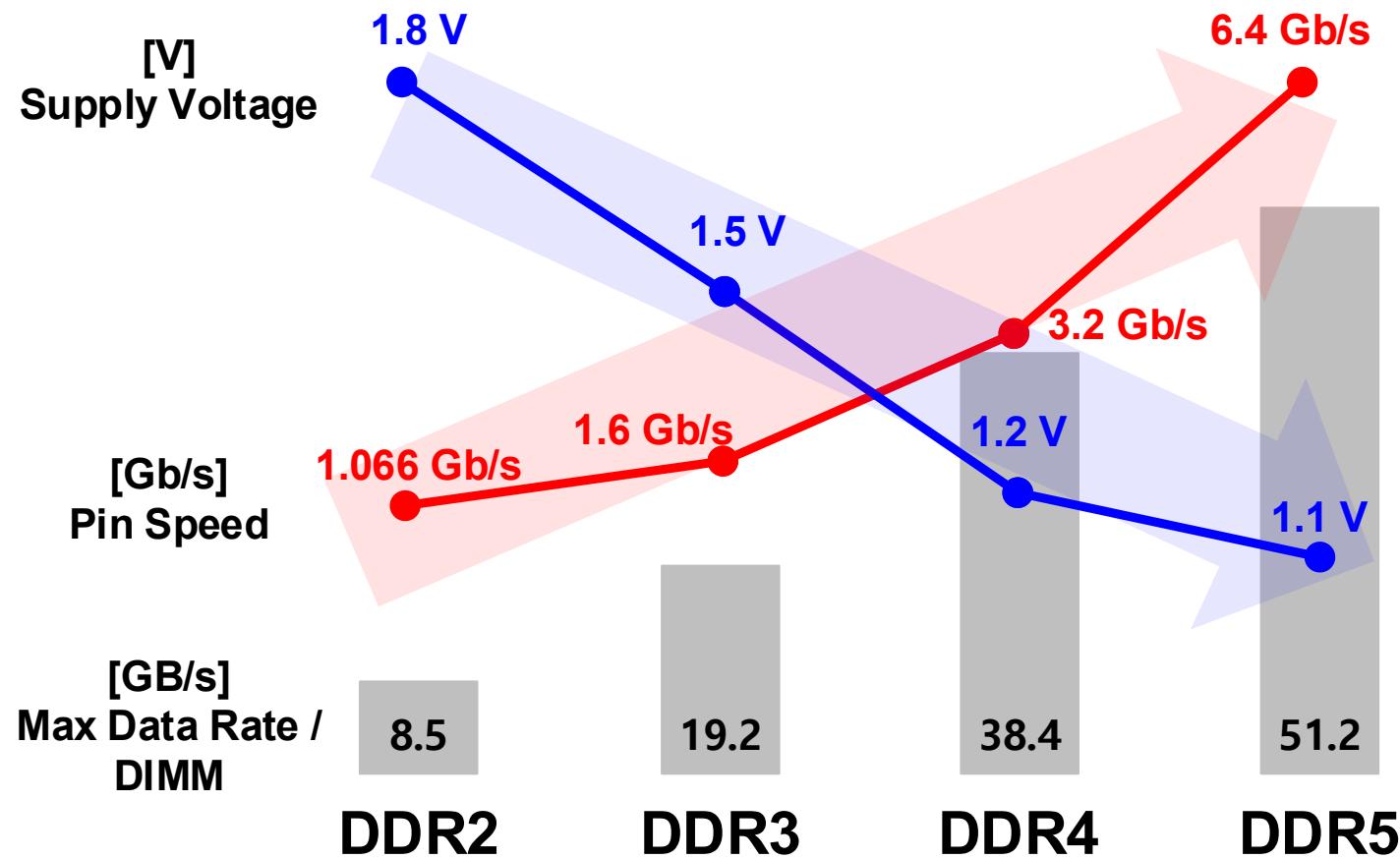
- GIO schemes for Low power consumption
- NBTI tolerable schemes for DLL
- Adaptive Body Bias for process variations
- Highly-accurate duty cycle detector
- High-speed schemes for transmitter

■ Implementation and measurements

■ Conclusion

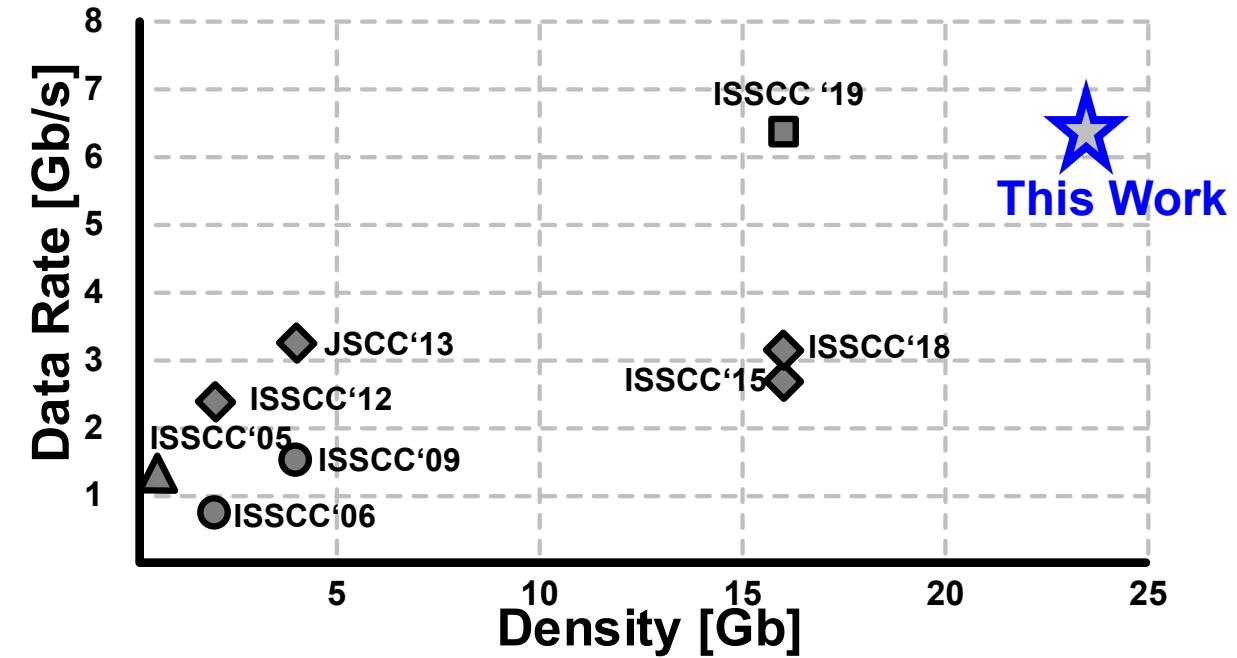
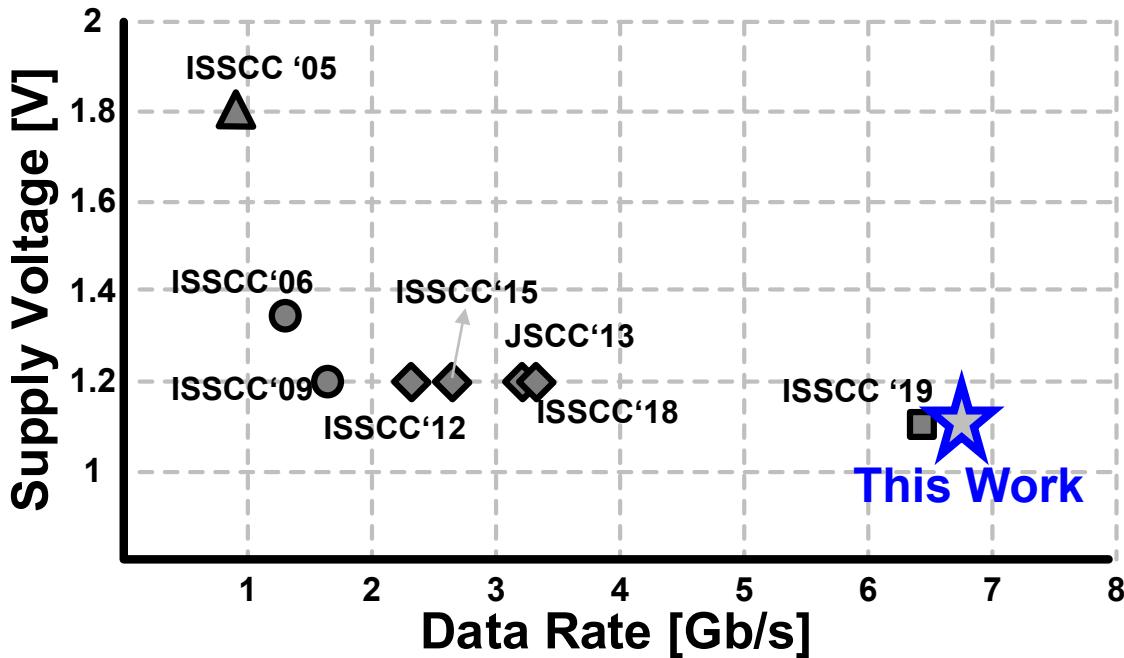
DDR5 Trend

- DRAM needs high-speed, low-power and high-density.



DDR5 Trend

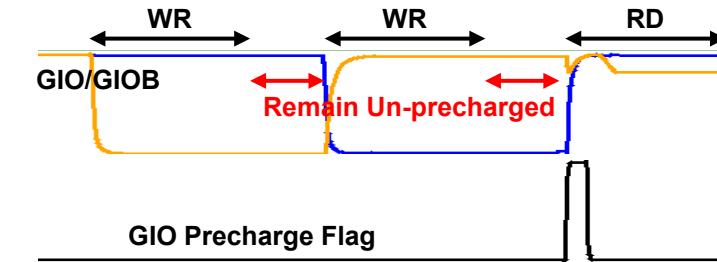
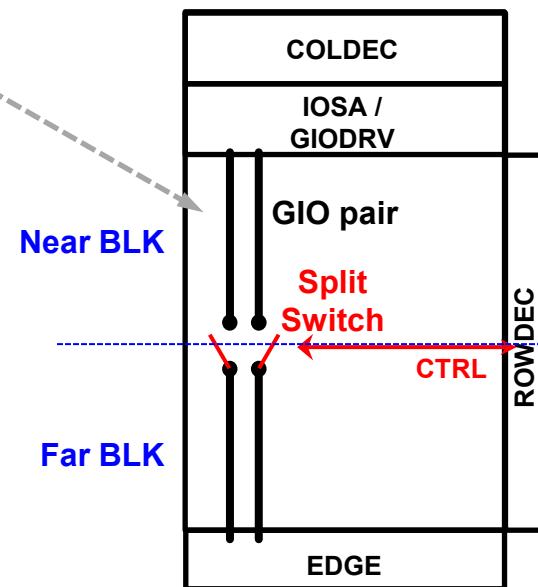
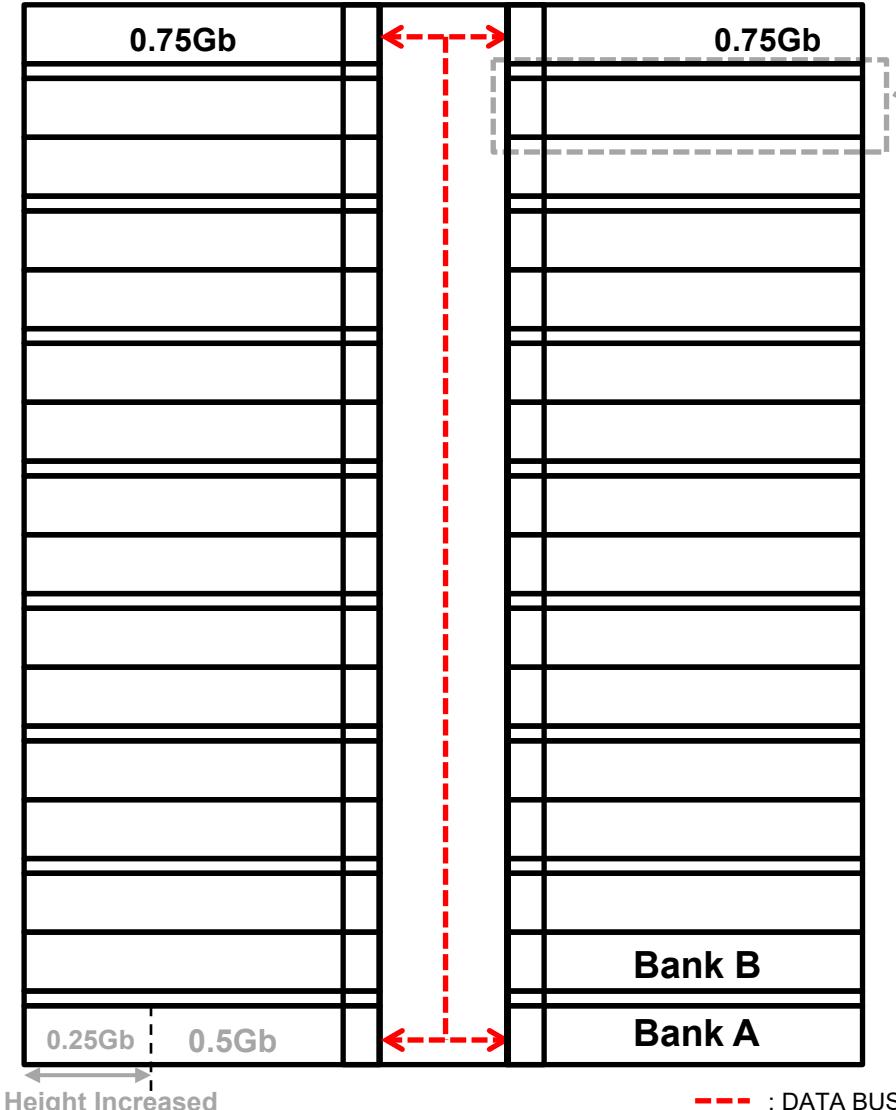
- **High-speed, high-density, and low-power consumption**
 - 24Gb DDR5 achieves 6.4 Gb/s/pin @ 1.1-V supply voltage



Outline

- **Introduction of DDR5**
- **Key schemes**
 - GIO schemes for Low power consumption
 - NBTI tolerable schemes for DLL
 - Adaptive Body Bias for process variations
 - Highly-accurate duty cycle detector
 - High-speed schemes for transmitter
- **Implementation and measurements**
- **Conclusion**

GIO schemes for Low power consumption



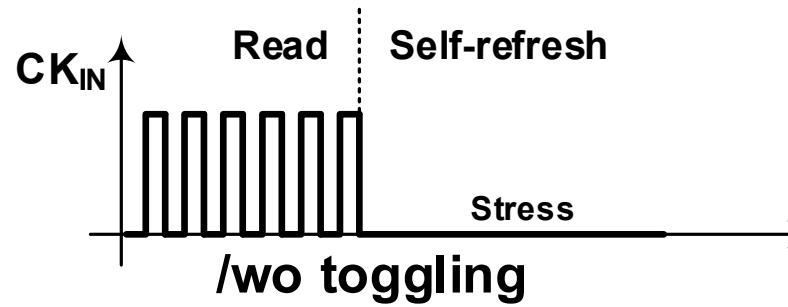
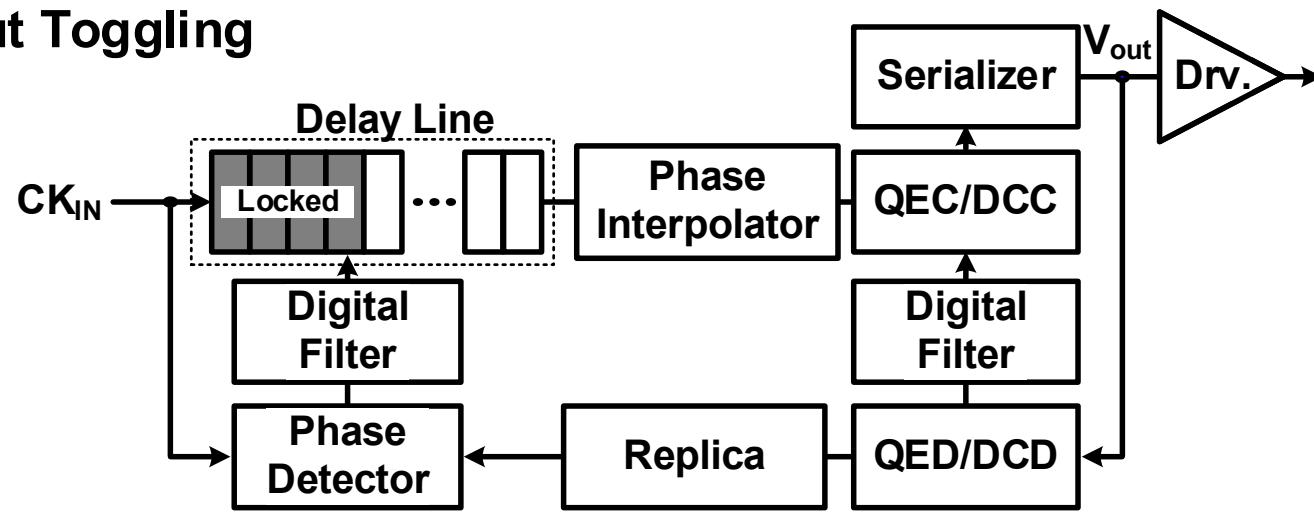
Product	DDR5	Proposed DDR5
Fabrication Process	10nm DRAM (2nd Generation)	10nm DRAM (4th Generation)
Die Density	16Gb	24Gb
Density per Bank	0.5Gb	0.75Gb
GIO Length	1x	1.5x
IDD4W Current (50% toggle case)	1x	0.93x

Outline

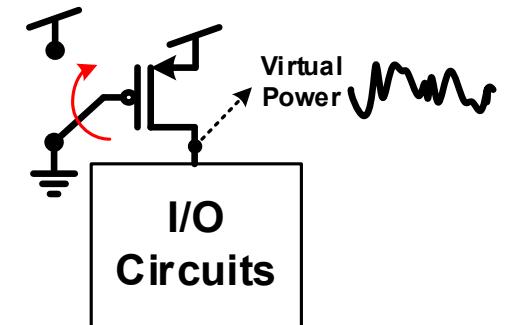
- **Introduction of DDR5**
- **Key schemes**
 - GIO schemes for Low power consumption
 - NBTI tolerable schemes for DLL
 - Adaptive Body Bias for process variations
 - Highly-accurate duty cycle corrector
 - High-speed schemes for transmitter
- **Implementation and measurements**
- **Conclusion**

NBTI tolerable schemes for DLL

Without Toggling



- **NBTI degradation**
 - Output performance degradation
 - V_{th} var. → duty/ delay var. ↑
- **Header only power gating**
 - Virtual power supply level ↓
 - Unwanted duty / delay var.

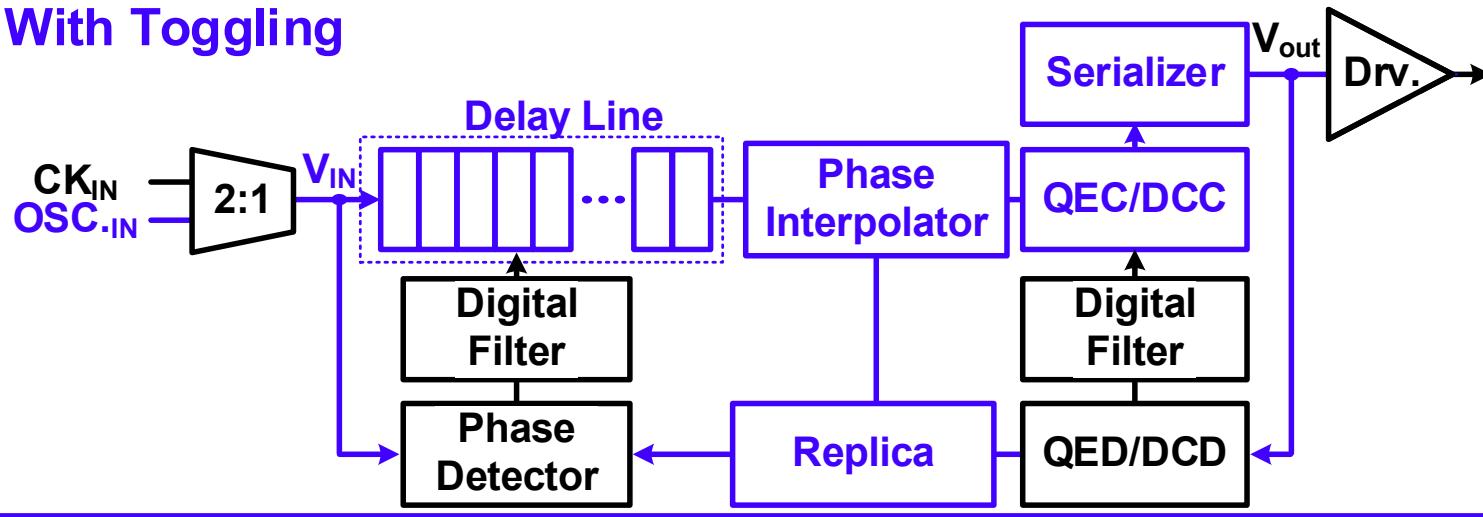


Header only power gating*

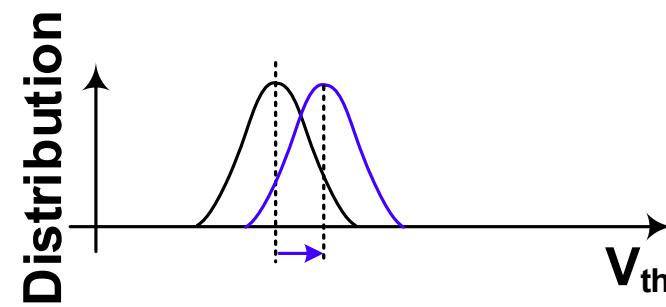
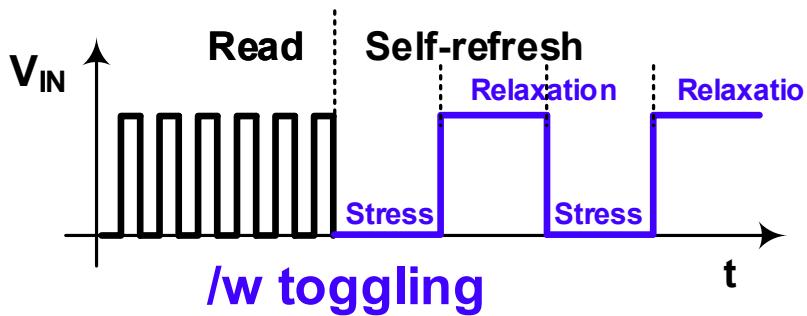
*REF: ISSCC 2018

NBTI tolerable schemes for DLL

With Toggling

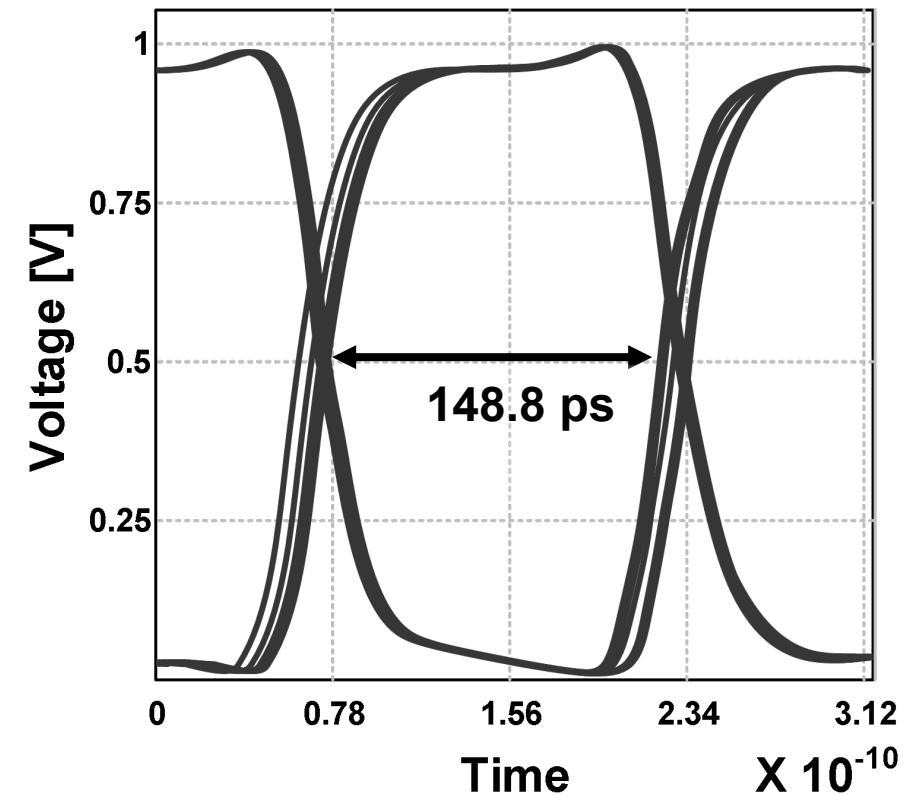
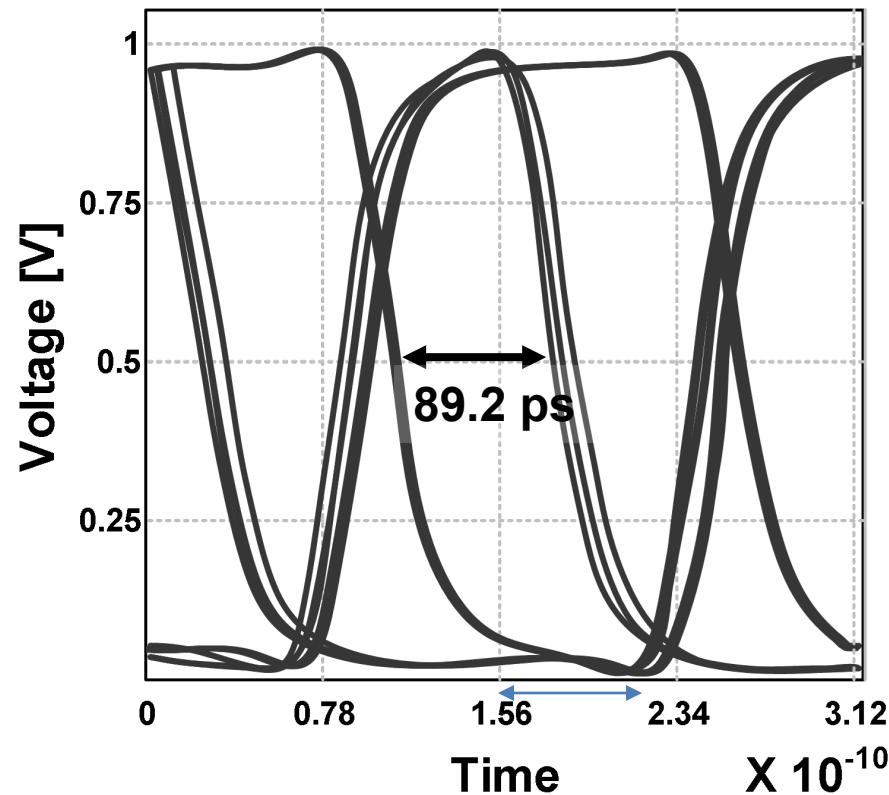


- **Toggling scheme**
 - Low freq. OSC in
 - Toggling all delay cells
- **Decreasing V_{th} var.**



NBTI tolerable schemes for DLL

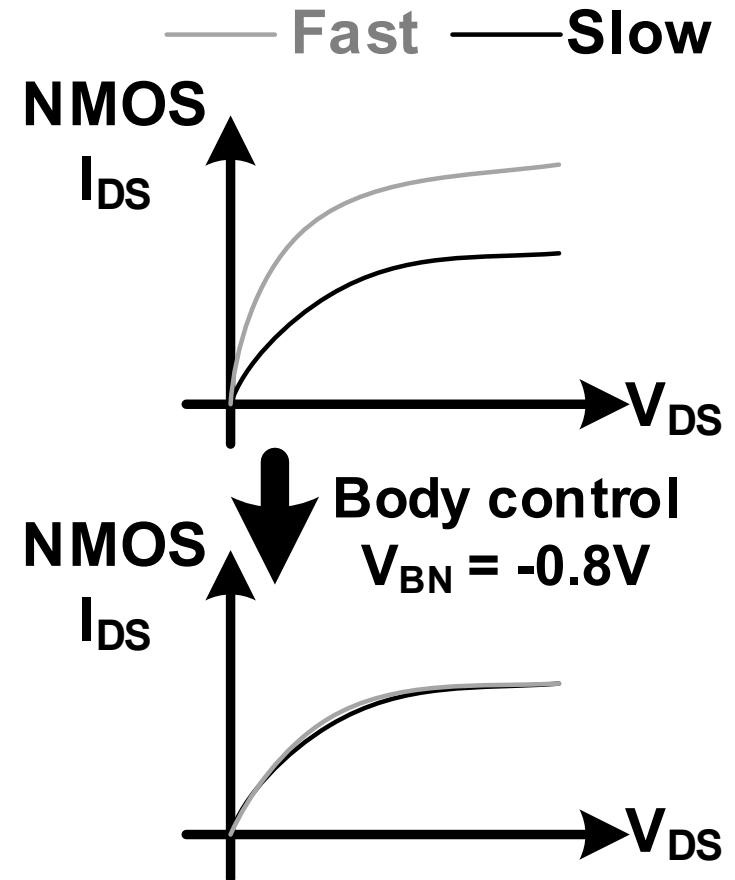
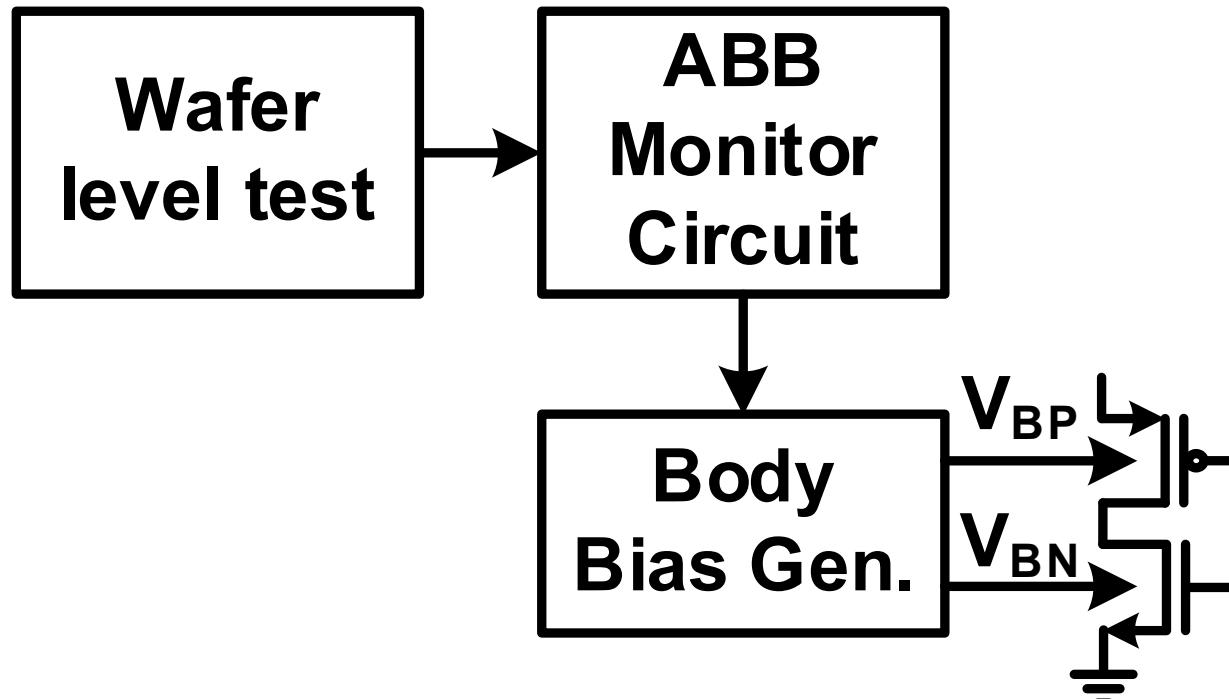
- Better performance with NBTI degradation
- Eye width: $89.2 \text{ ps} \rightarrow 148.8 \text{ ps}$ ($1\text{UI} = 156.25\text{ps}$) : 66% ▲



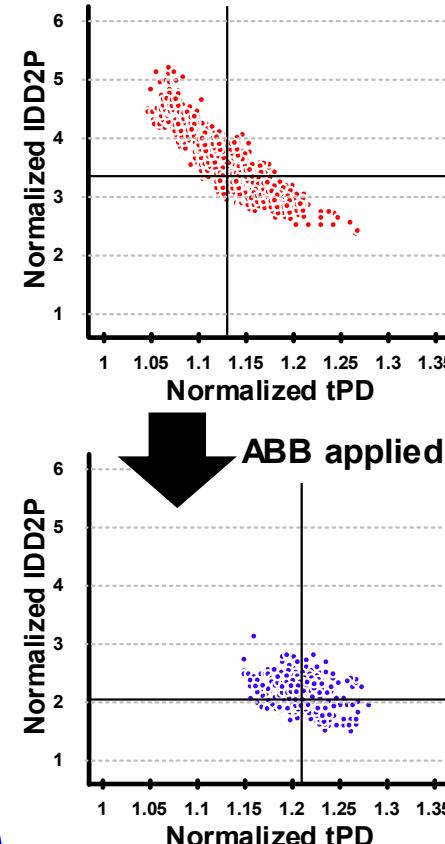
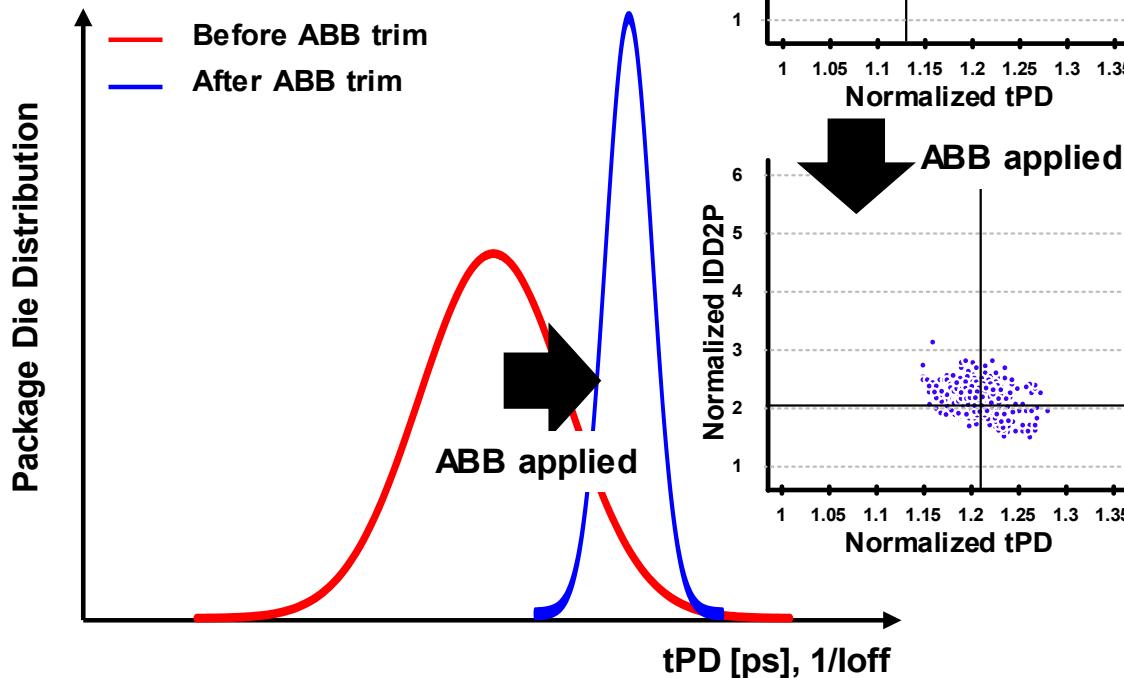
Outline

- **Introduction of DDR5**
- **Key schemes**
 - GIO schemes for Low power consumption
 - NBTI tolerable schemes for DLL
 - Adaptive Body Bias for process variations
 - Highly-accurate duty cycle detector
 - High-speed schemes for transmitter
- **Implementation and measurements**
- **Conclusion**

Adaptive Body Bias for process variations

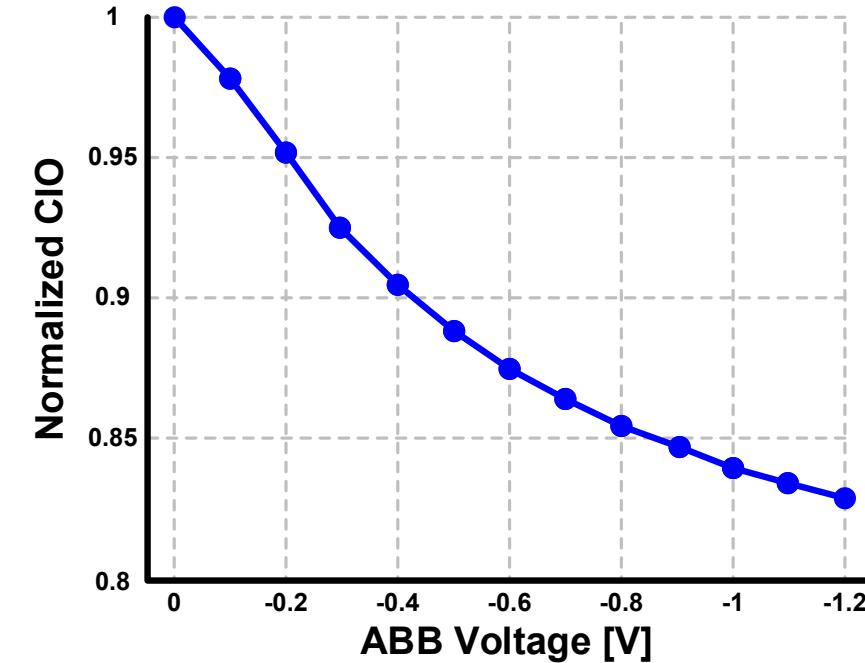


Adaptive Body Bias for process variations



■ Applied ABB Voltage

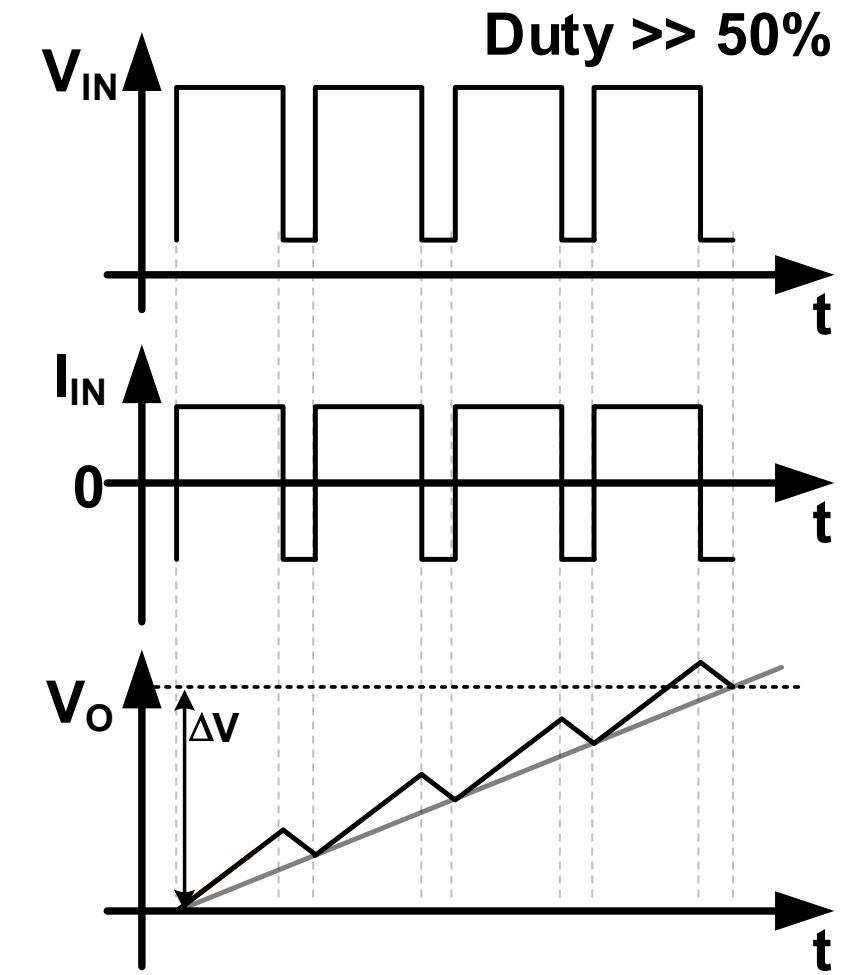
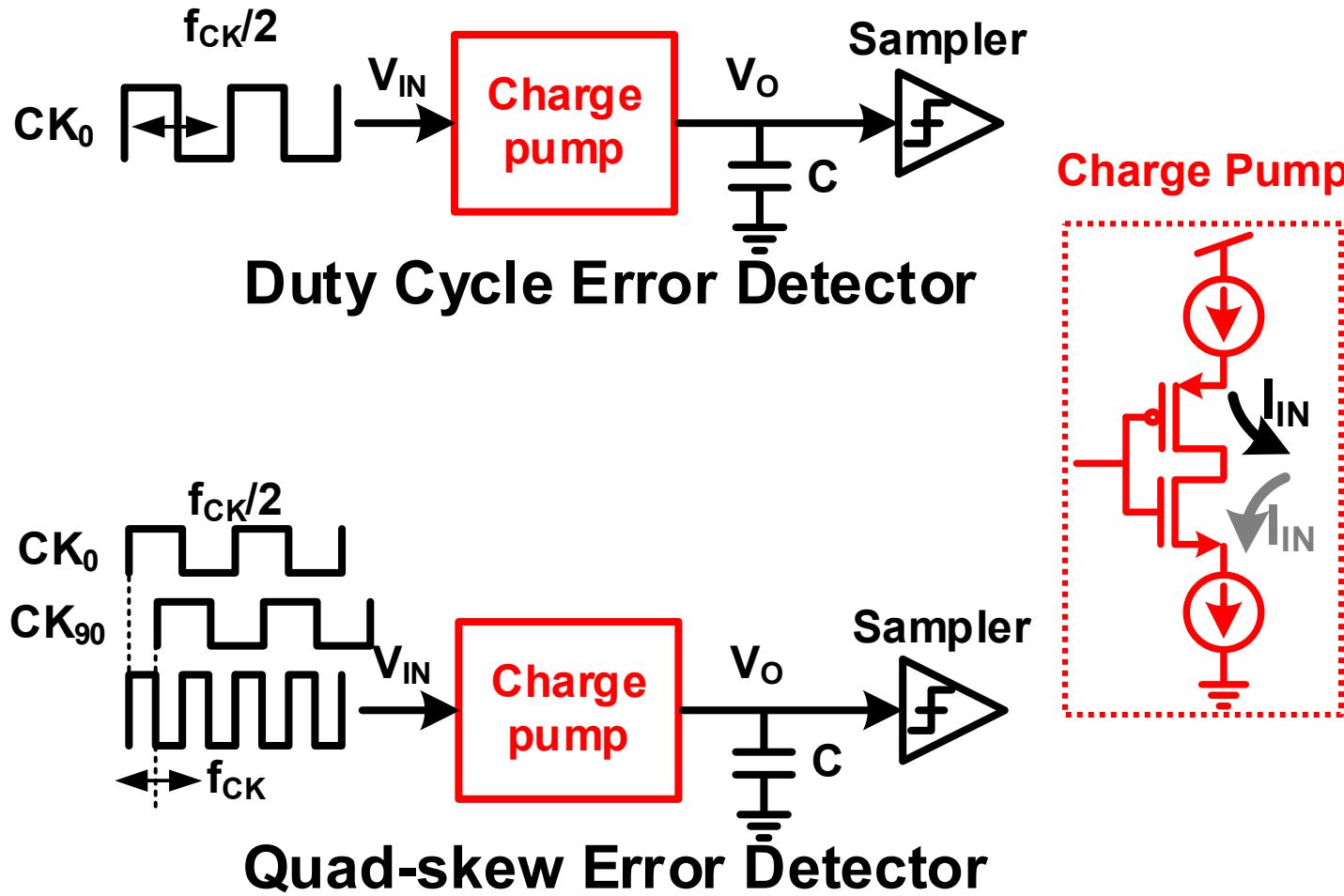
- Distribution of tPD & IDD2P ↓
- Normalized CIO: - 15% ☺
→ Decreasing chip size / power consumption



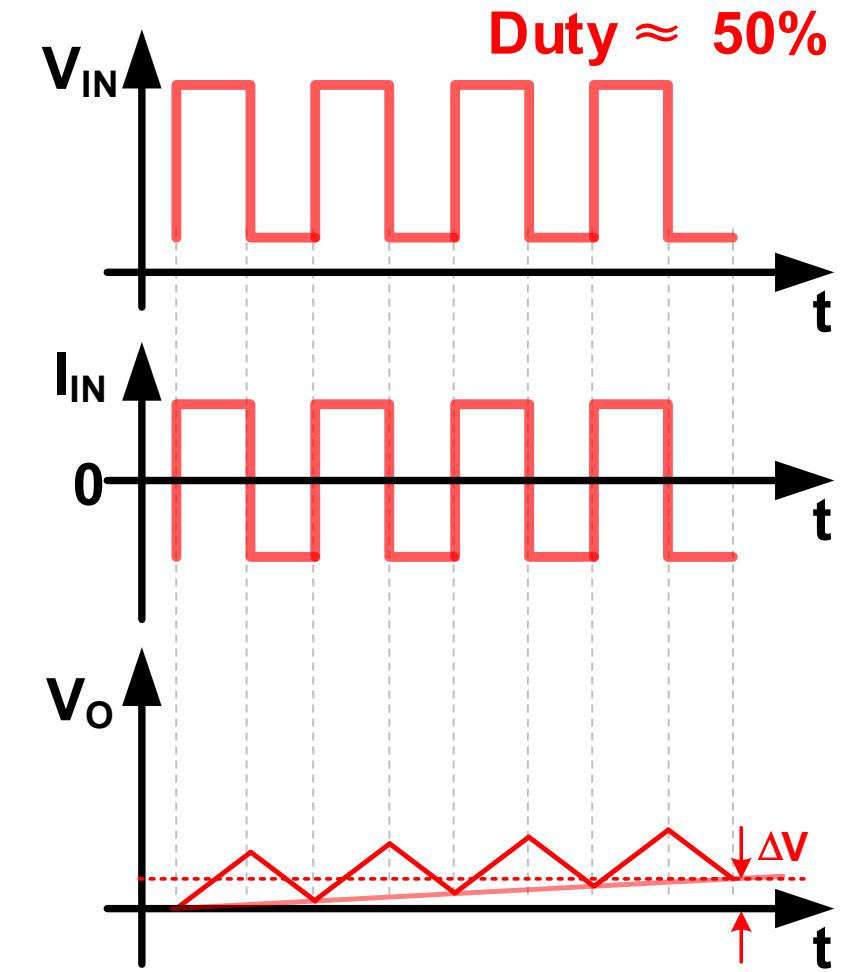
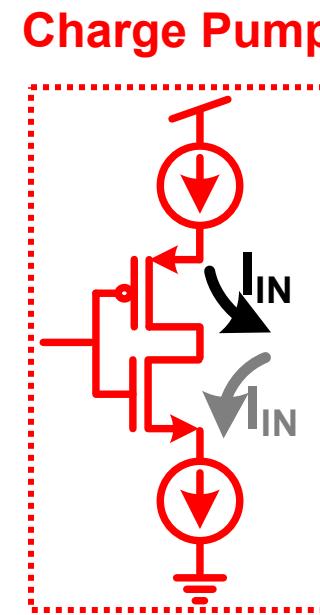
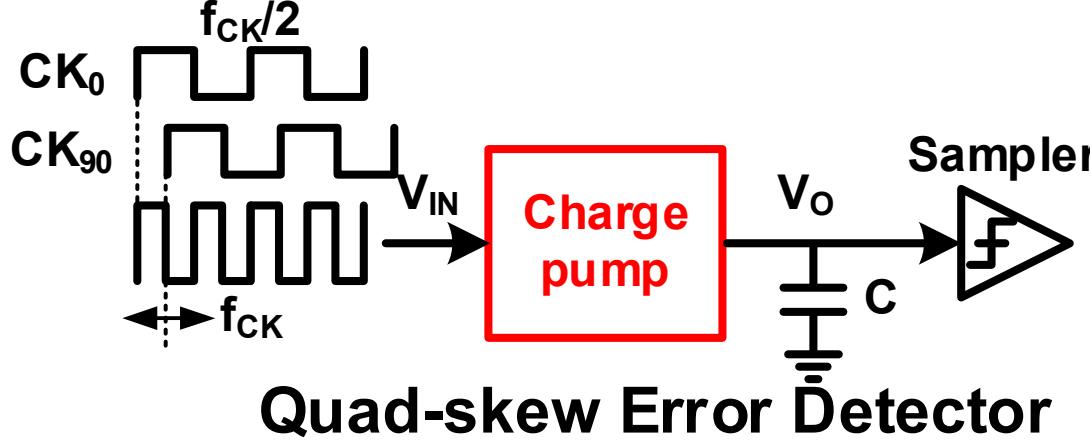
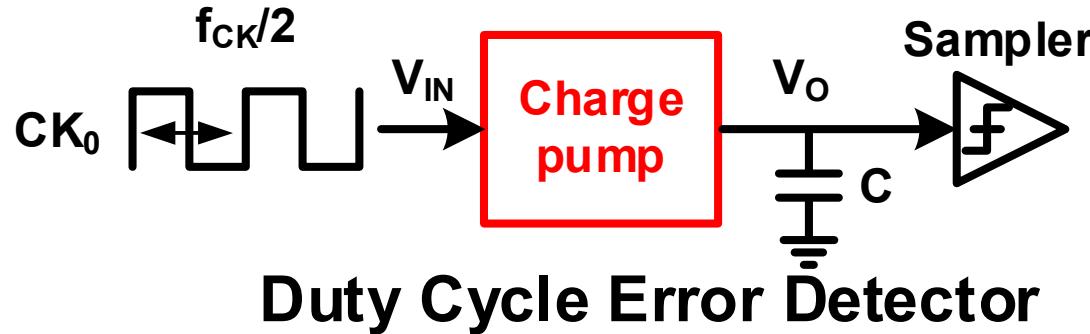
Outline

- **Introduction of DDR5**
- **Key schemes**
 - GIO schemes for Low power consumption
 - NBTI tolerable schemes for DLL
 - Adaptive Body Bias for process variations
 - **Highly-accurate duty cycle detector**
 - High-speed schemes for transmitter
- **Implementation and measurements**
- **Conclusion**

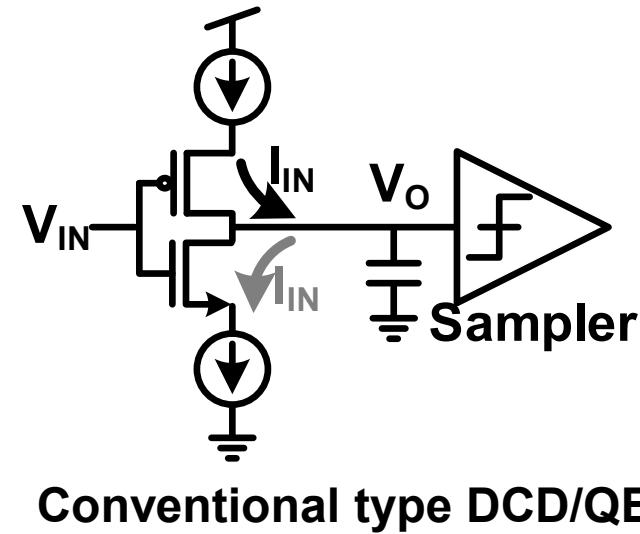
Highly-accurate duty cycle detector



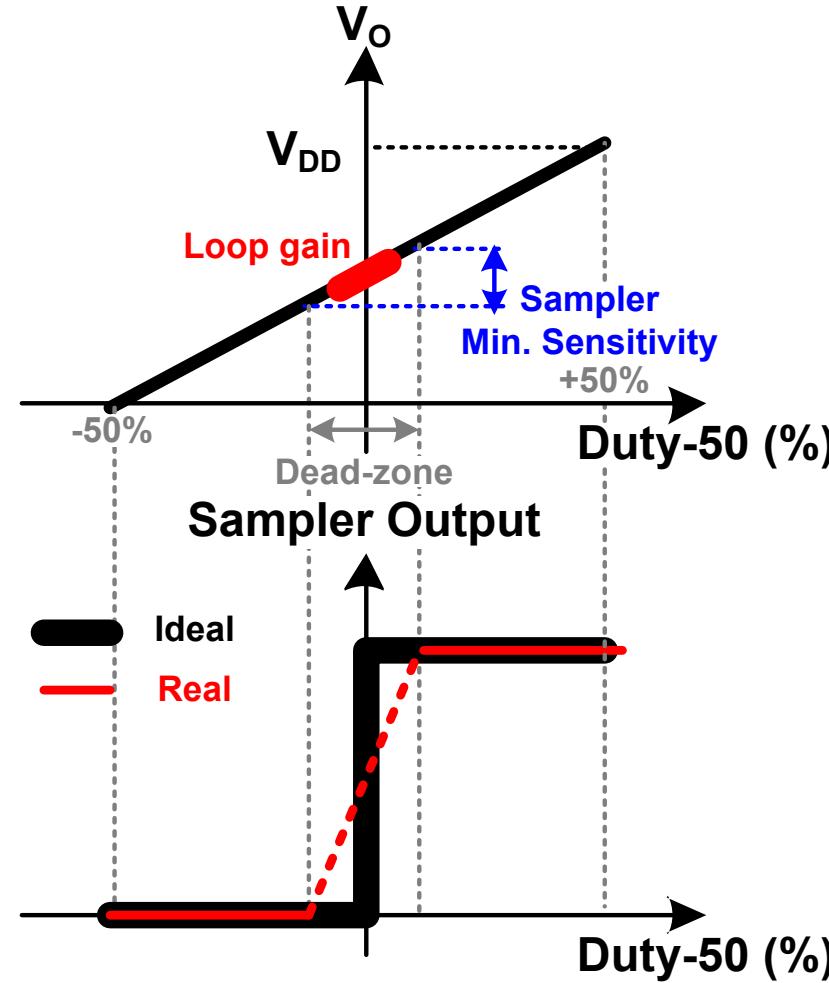
Highly-accurate duty cycle detector



Highly-accurate duty cycle detector

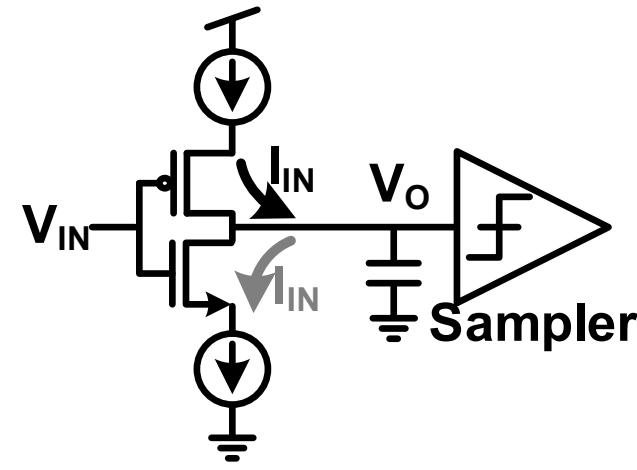


Conventional type DCD/QED

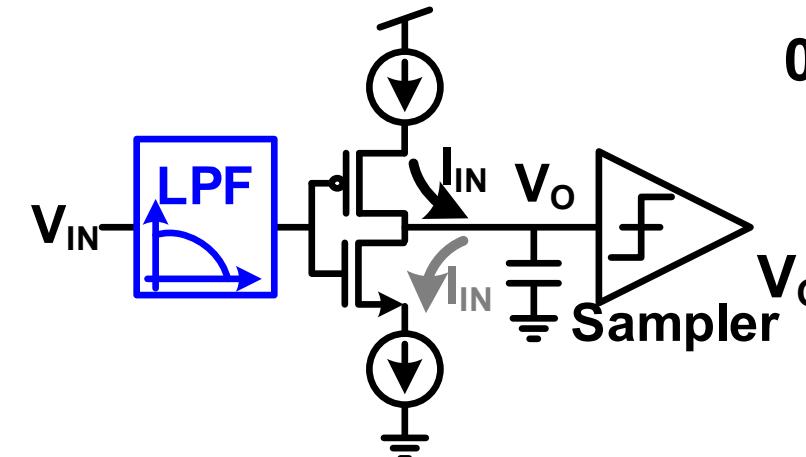


- **Loop Gain**
 - Slope of charge pump output
- **Dead-zone**
 - In real, sampler has min. sensitivity
 - Cannot sampling perfectly
- **Previous type DCD/ QED**
 - Sampler sensitivity : Power consumption ☹
 - Charge pump current↑ : Power consumption ☹
Update period↑ : Loop bandwidth↓ ☹

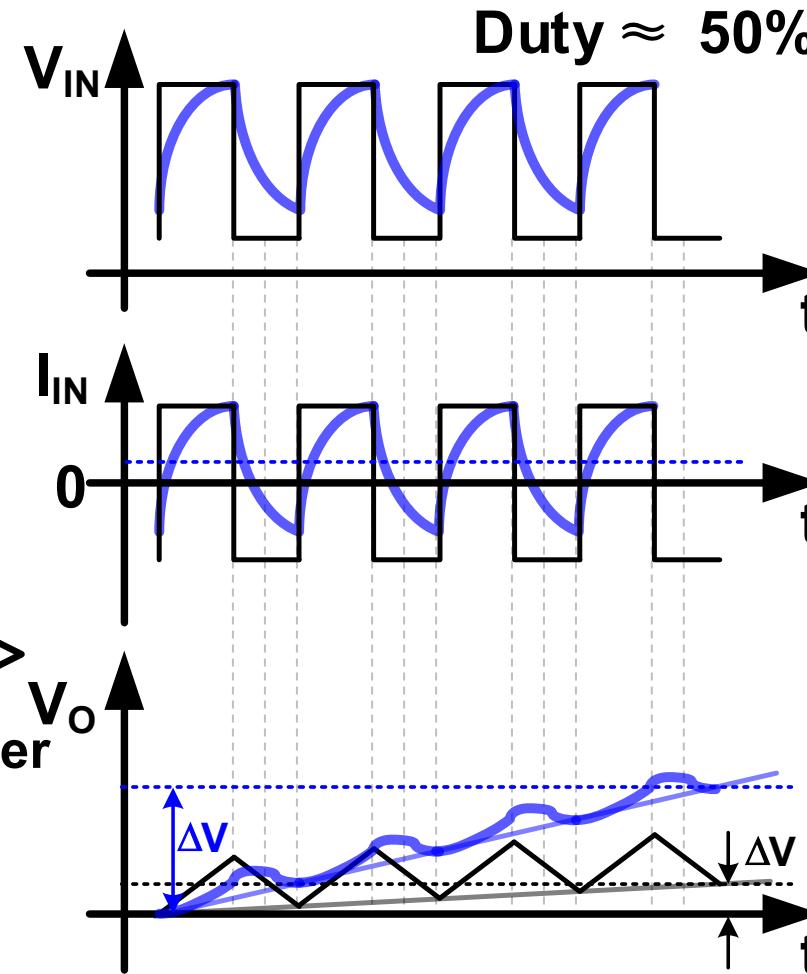
Highly-accurate duty cycle detector



Conventional type DCD/QED

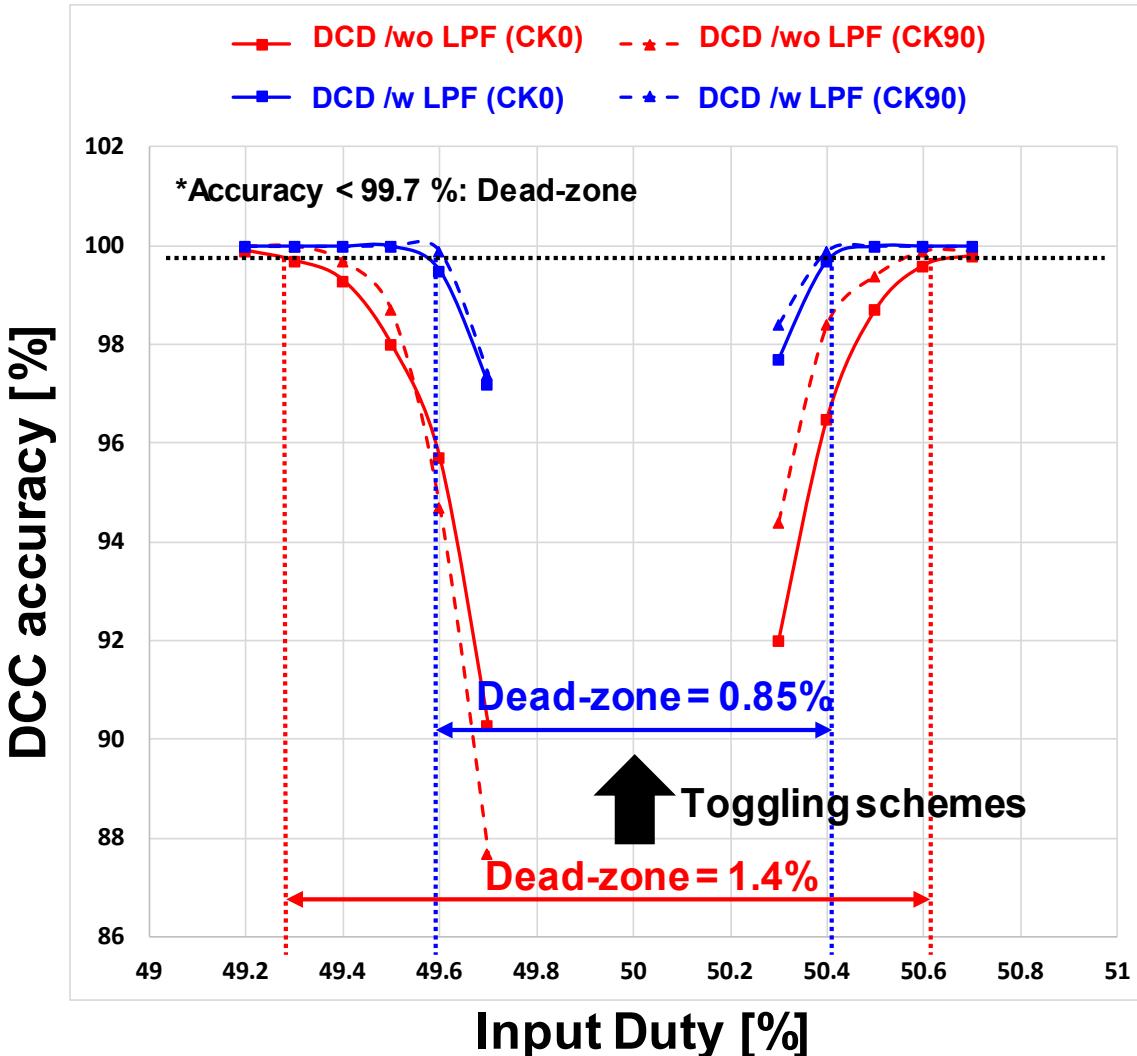


Proposed type DCD/QED



- Much higher loop gain
 - DC current + AC current
 - DC : Common-mode voltage
 - AC : Input duty cycle
- Integrate charge faster

Highly-accurate duty cycle detector



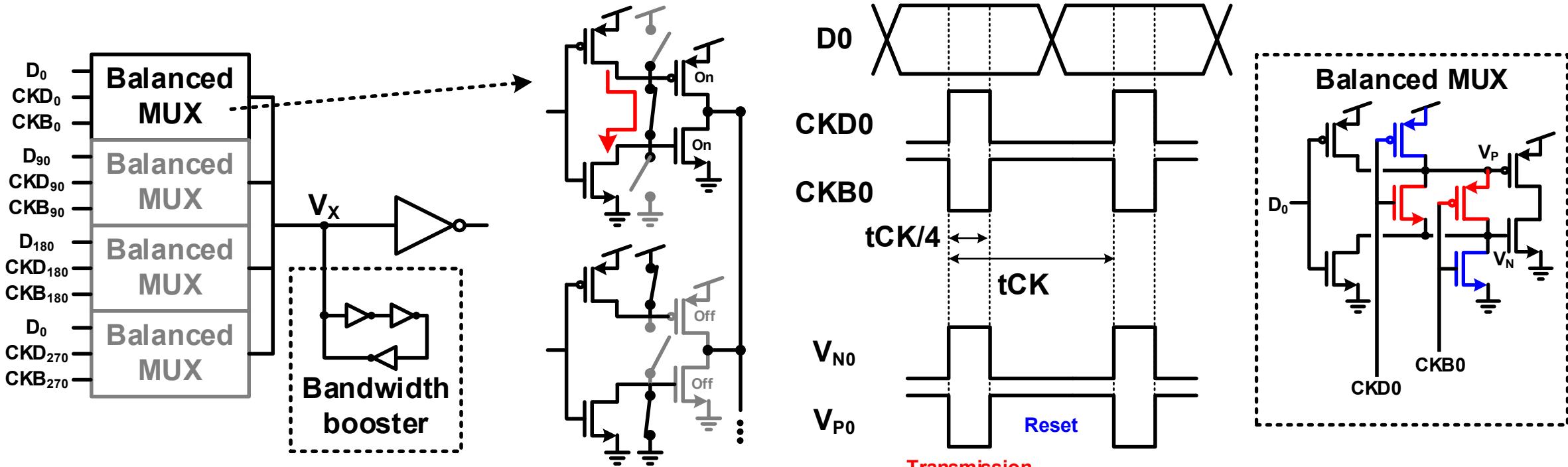
■ Monte-carlo simulation

- Under 99.7 % accuracy → Dead-zone
- Carefully design : $BW_{CP} /w LPF >$ Loop bandwidth

Outline

- **Introduction of DDR5**
- **Key schemes**
 - GIO schemes for Low power consumption
 - NBTI tolerable schemes for DLL
 - Adaptive Body Bias for process variations
 - Highly-accurate duty cycle detector
 - High-speed schemes for transmitter
- **Implementation and measurements**
- **Conclusion**

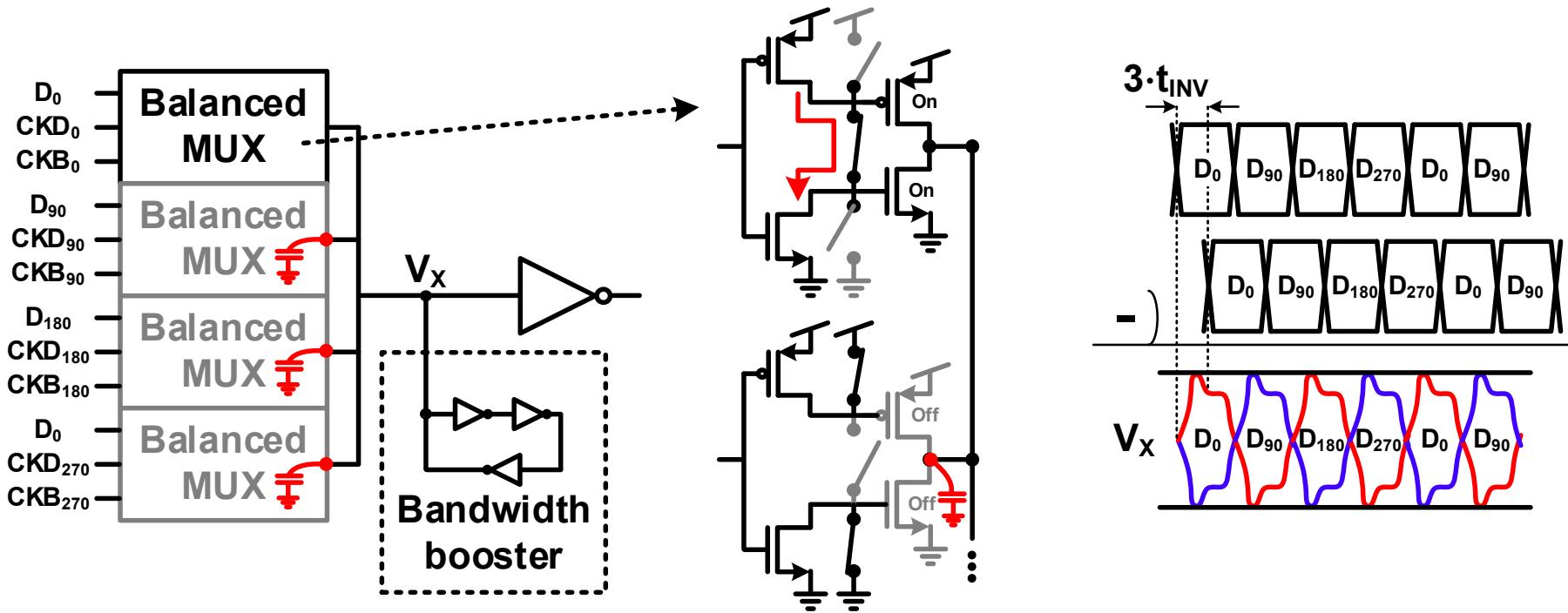
High-speed schemes for transmitter



Balanced MUX (4:1 MUX)

- Reset gates for output “off”, transmission gates for input sampling
- No NAND/NOR : No data dependent jitter by coupling, active area ↓ ☺
- Pre-DRV. : Power reduction ☺
- Output loading cap. ↑ ☹

High-speed schemes for transmitter



■ Bandwidth booster

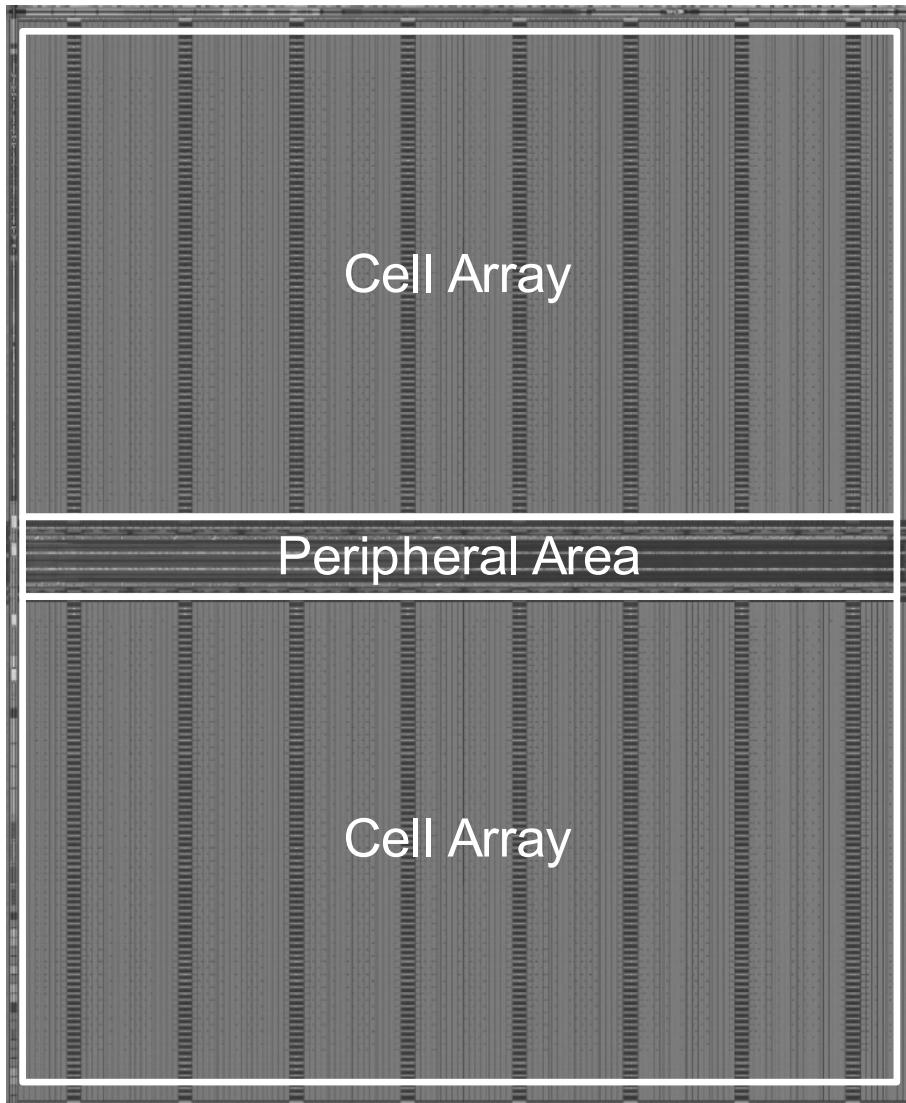
- Equalizing internal ISI ☺ from balanced MUXs (loading cap. \uparrow)
- Simply designed
- $3 \cdot t_{INV} > 1\text{UI} \rightarrow$ Signal distorted ☹

Outline

- **Introduction of DDR5**
- **Key schemes**
 - GIO schemes for Low power consumption
 - NBTI tolerable schemes for DLL
 - Adaptive Body Bias for process variations
 - Highly-accurate duty cycle detector
 - High-speed schemes for transmitter
- **Implementation and measurements**
- **Conclusion**

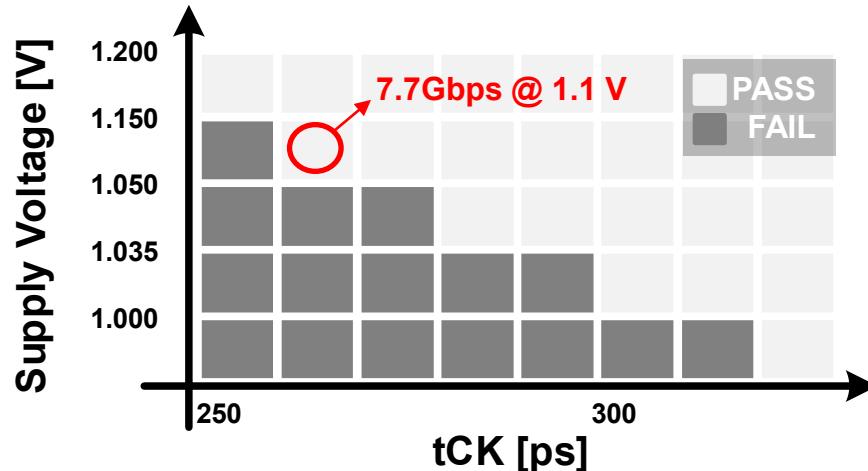
Chip Implementation

- **Process : 10 nm DRAM
(4th generation)**
- **Density : 24Gb/CH**
- **Max data rate : 6.4Gb/s/pin**
- **Supply voltage**
 - ✓ **VDDQ / VDD = 1.1V / 1.1V**

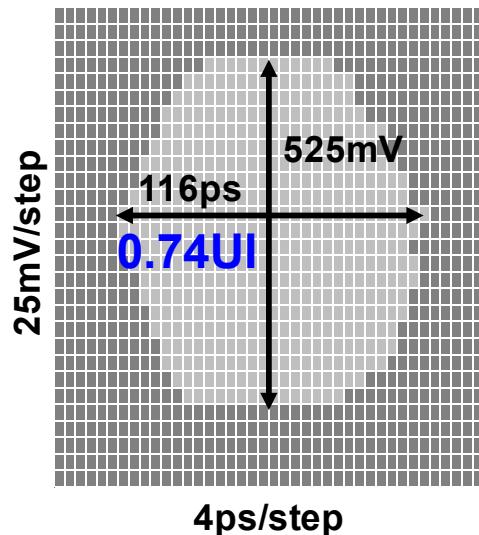


28.7: A 1.1-V 6.4-Gb/s/pin 24-Gb DDR5 SDRAM with a Highly-Accurate Duty Corrector and NBTI-Tolerant DLL

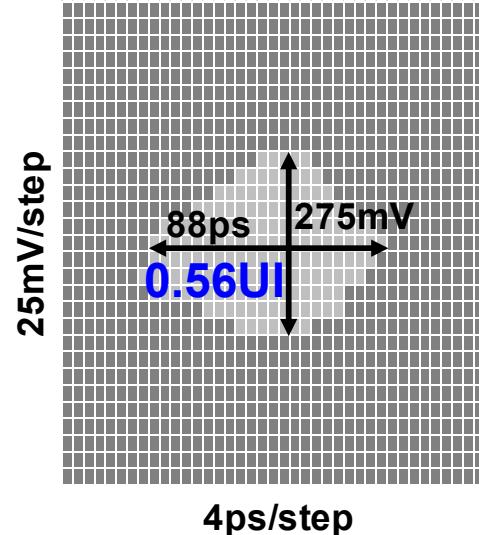
Measurements



Read shmoo @ 6.4Gbps



Write shmoo @ 6.4Gbps



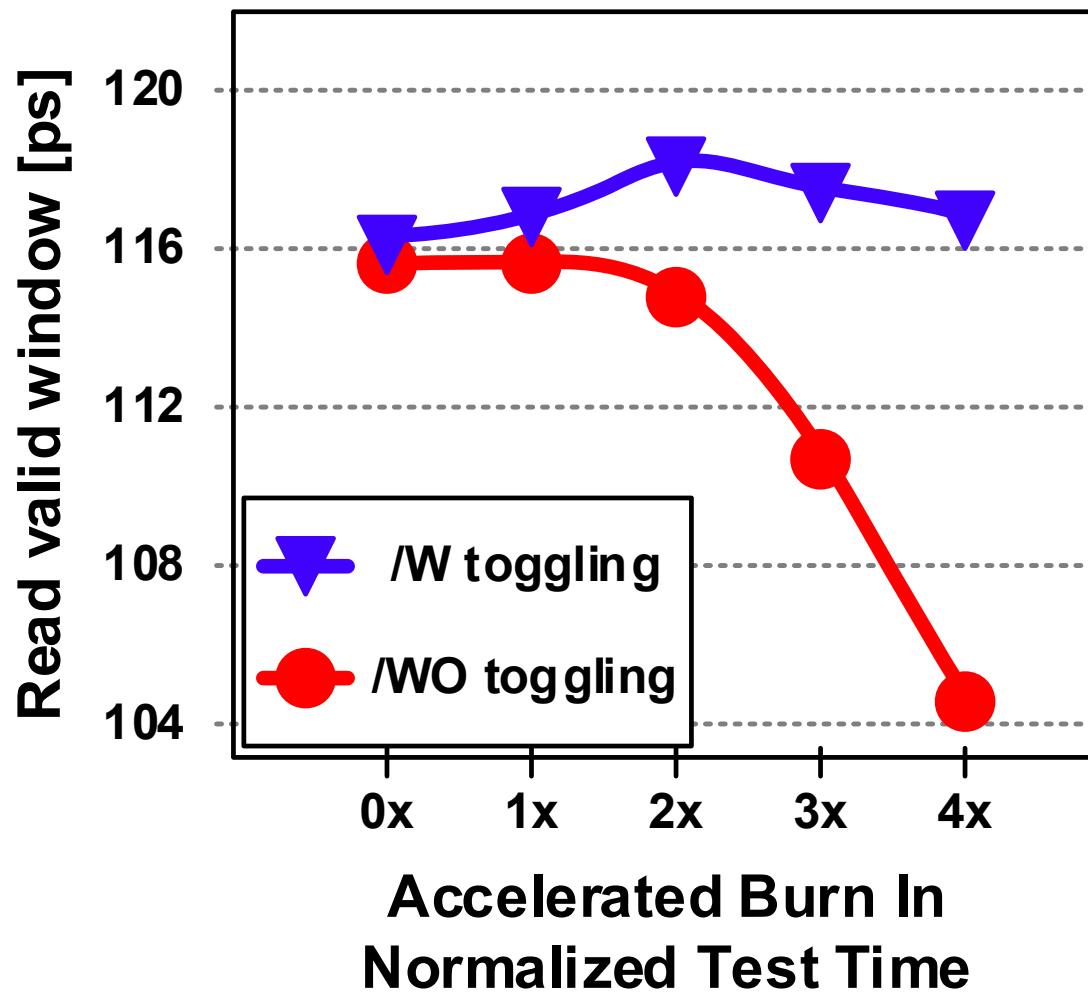
■ Frequency-voltage shmoo

- Max freq. 7.7Gb/s @ 1.1V

■ Read/ Write shmoo @ 6.4Gbps

- Valid window = 116 ps @ RD (0.74UI)
- Valid window = 88 ps @ WR (0.56UI)
 - No Eq. (CTLE, DFE)
 - Low power mode

Measurements



- **Accelerated Burn In test**
 - Valid window = 116 ps @ RD

Outline

- **Introduction of DDR5**
- **Key schemes**
 - GIO schemes for Low power consumption
 - NBTI tolerable schemes for DLL
 - Adaptive Body Bias for process variations
 - Highly-accurate duty cycle corrector
 - High-speed schemes for transmitter
- **Implementation and measurements**
- **Conclusion**

Conclusion

- **6.4 Gb/s/pin 24Gb DDR5 SDRAM is implemented**
- **High-density, High-speed and Low-power**
 - GIO schemes for Low power consumption
 - IDD 4W: ▼38%
(Normalized power consumption with same density and same fabrication)
 - NBTI tolerable schemes for DLL
 - Accelerated burn-in test: Eye width ▲12%
 - Reverse Adaptive Body Bias
 - Decreasing process variation
 - Highly-accurate duty cycle error detector
 - Higher loop gain ▲60%
 - High-speed schemes for transmitter
 - Overcome the disadvantages of ABB aligning all corners to slow side

Thank you

A 1.1-V 16-Gb DDR5 DRAM with Probabilistic-Aggressor Tracking, Refresh-Management Functionality, Per-Row Hammer Tracking, a Multi-Step Precharge, and Core-Bias Modulation for Security and Reliability Enhancement

Woongrae Kim, Chulmoon Jung, Seongnyuh Yoo, Duckhwa Hong, Jeongjin Hwang, Jungmin Yoon, Ohyong Jung, Joonwoo Choi, Sanga Hyun, Mankeun Kang, Sangho Lee, Dohong Kim, Sanghyun Ku, Donhyun Choi, Nogeun Joo, Sangwoo Yoon, Junseok Noh, Byeongyang Go, Cheolhoe Kim, Sunil Hwang, Mihyun Hwang, Seol-Min Yi, Hyungmin Kim, Sanghyuk Heo, Yeonsu Jang, Kyoungchul Jang, Shinho Chu, Yoonna Oh, Kwidong Kim, Junghyun Kim, Soohwan Kim, Jeongtae Hwang, Sangil Park, Junphyo Lee, Inchul Jeong, Joohwan Cho, Jonghwan Kim



SK Hynix



Self Introduction

■ First author name (speaker) : Woongrae Kim

■ Education

- B.S. in ECE from Hanyang University, Seoul, Korea in 2009
- M.S. in ECE from Georgia Institute of Technology, GA, USA in 2015
- Ph.D. in ECE from Georgia Institute of Technology, GA, USA in 2016



■ Experience

- Principal Engineer, DRAM Design Division, SK Hynix, Korea (2016~)

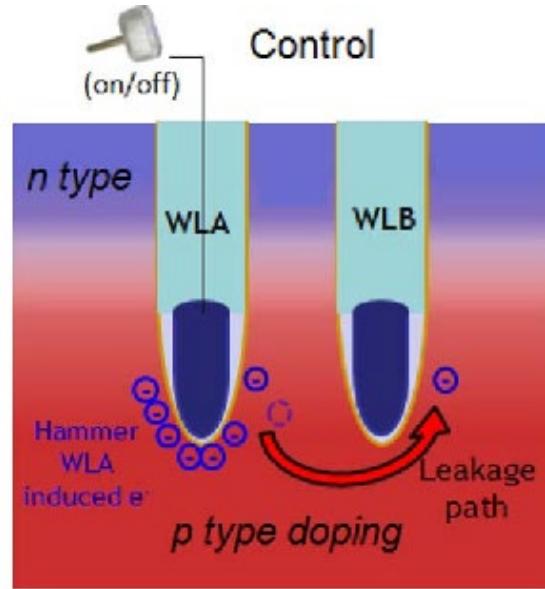
■ Research Interest

- Low power and high-speed circuit design
- Scalable memory systems
- DRAM security and reliability

Outline

- **Introduction**
- **Architecture Overview**
- **Key Schemes**
 - Probabilistic aggressor tracking (PAT)
 - Refresh management function (RFM)
 - Per-row hammer tracking (PRHT)
 - Multi-step precharge circuit
 - Core-bias modulation
- **Measurement Results and Chip Implementation**
- **Conclusion**

Introduction



Intrinsic row hammer tolerance
= f(amount of charge in cell, e-field)

[1] Suppression of Row Hammer Effect by Doping Profile Modification in Saddle-Fin Array Devices for Sub-30-nm DRAM technology, Chia-Ming Yang

- DRAM manufacturers are facing technology scaling challenges due to row hammer and refresh retention time beyond 1a-nm

Introduction

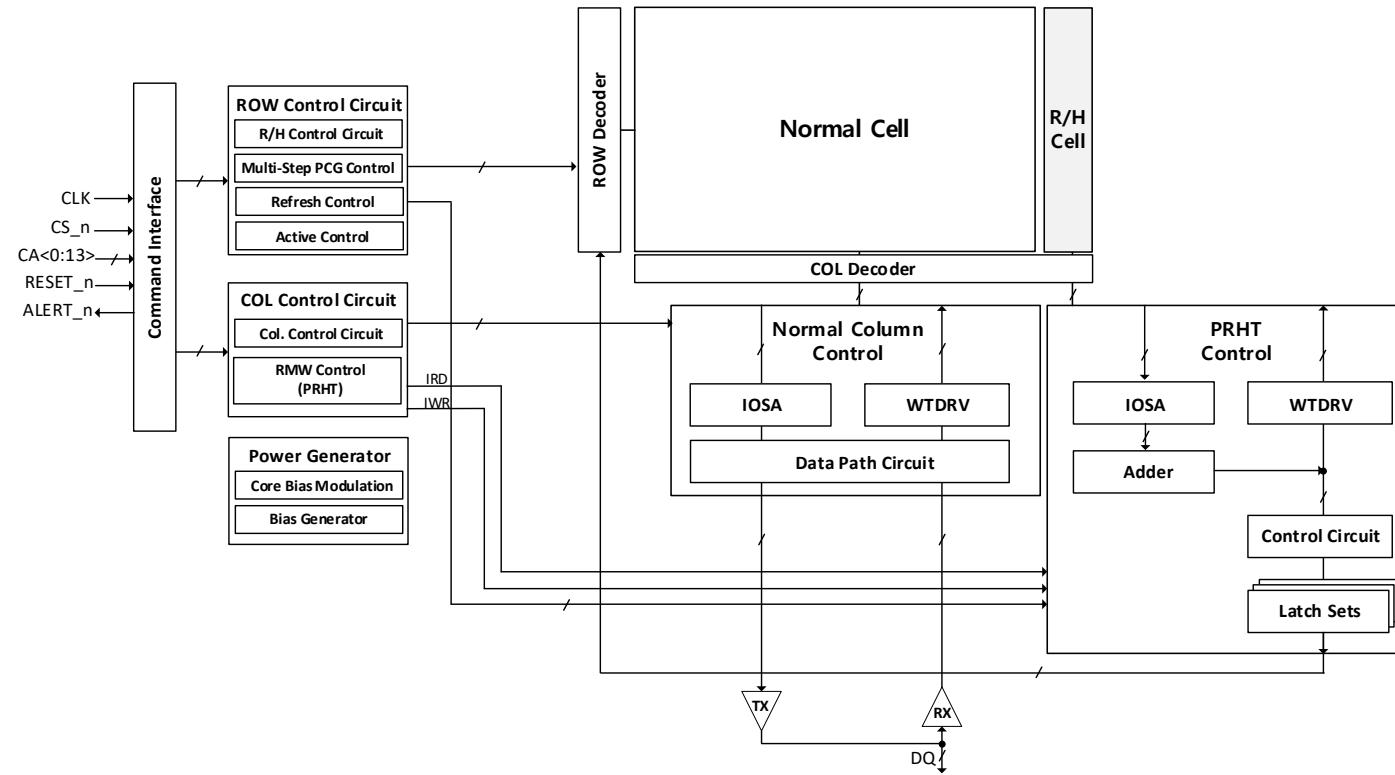
■ How to protect DRAM from row hammer

- Activate adjacent rows with carefully sampled active addresses
- Execute additional refresh commands when malicious attacks are detected
- Reinforce intrinsic row-hammer tolerance

■ How to maximize refresh retention time

- Control cell transistors to find the optimal tradeoff between intrinsic row-hammer tolerance and refresh retention time

Architecture Overview of 1anm 16Gb DDR5

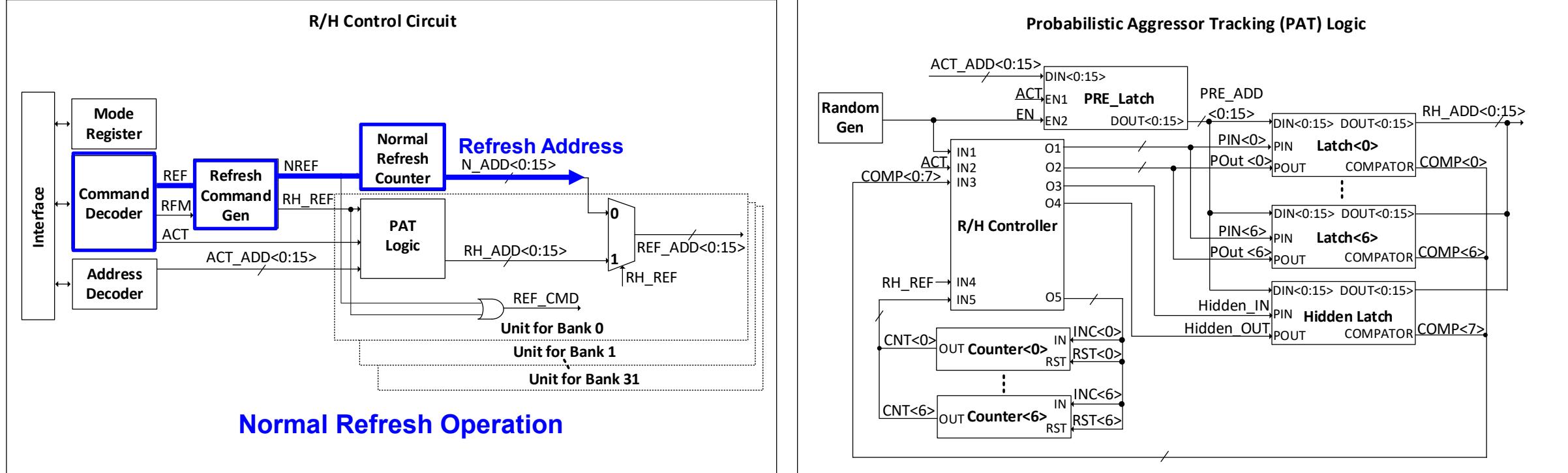


- Probabilistic aggressor tracking, per-row hammer tracking
- Refresh management function (RFM)
- Multi-step precharge circuit, core bias modulation scheme

Outline

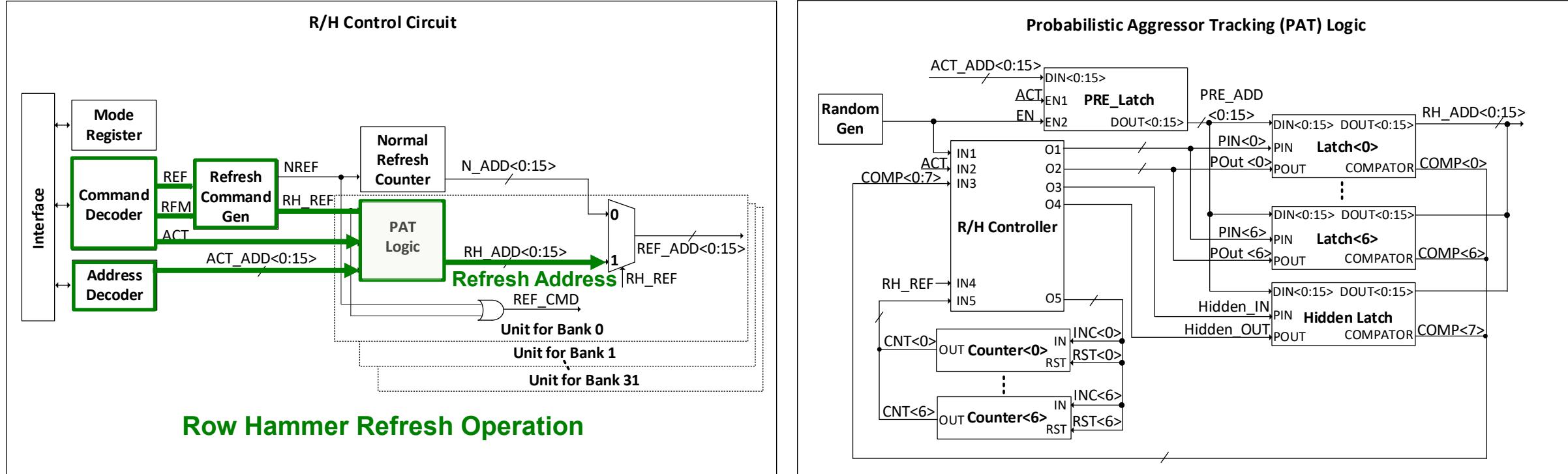
- **Introduction**
- **Architecture Overview**
- **Key Schemes**
 - Probabilistic aggressor tracking (PAT)
 - Refresh management function (RFM)
 - Per-row hammer tracking (PRHT)
 - Multi-step precharge circuit
 - Core-bias modulation
- **Measurement Results and Chip Implementation**
- **Conclusion**

Probabilistic Aggressor Tracking (PAT)



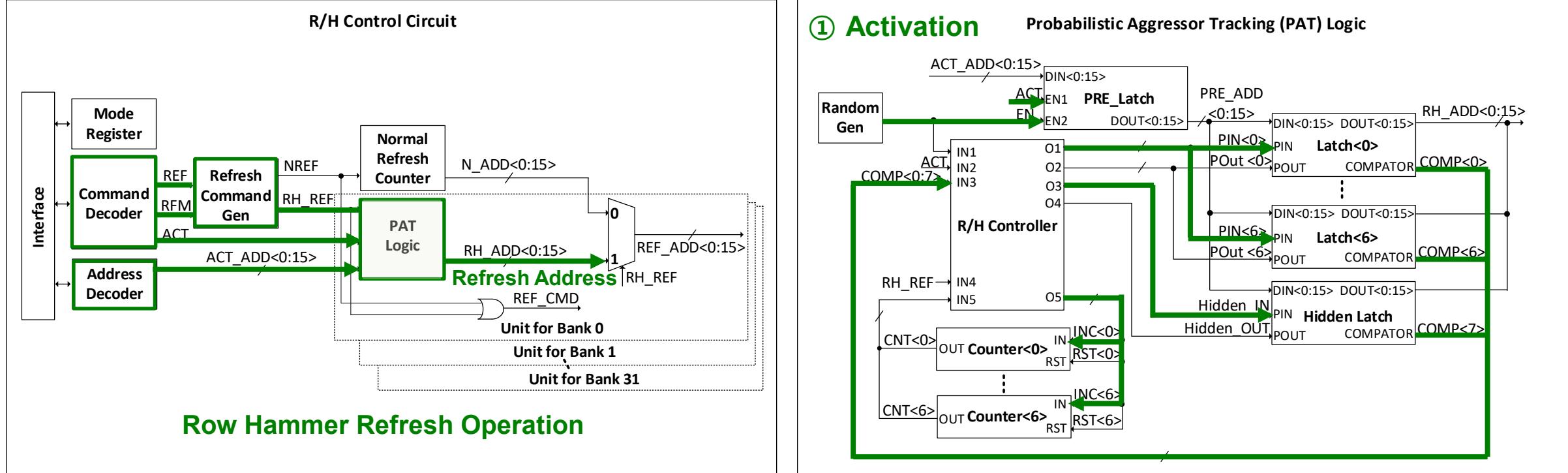
- Probabilistic approach to improve aggressor tracking accuracy
- Cost-effective since it is implemented within the peripheral area

Probabilistic Aggressor Tracking (PAT)



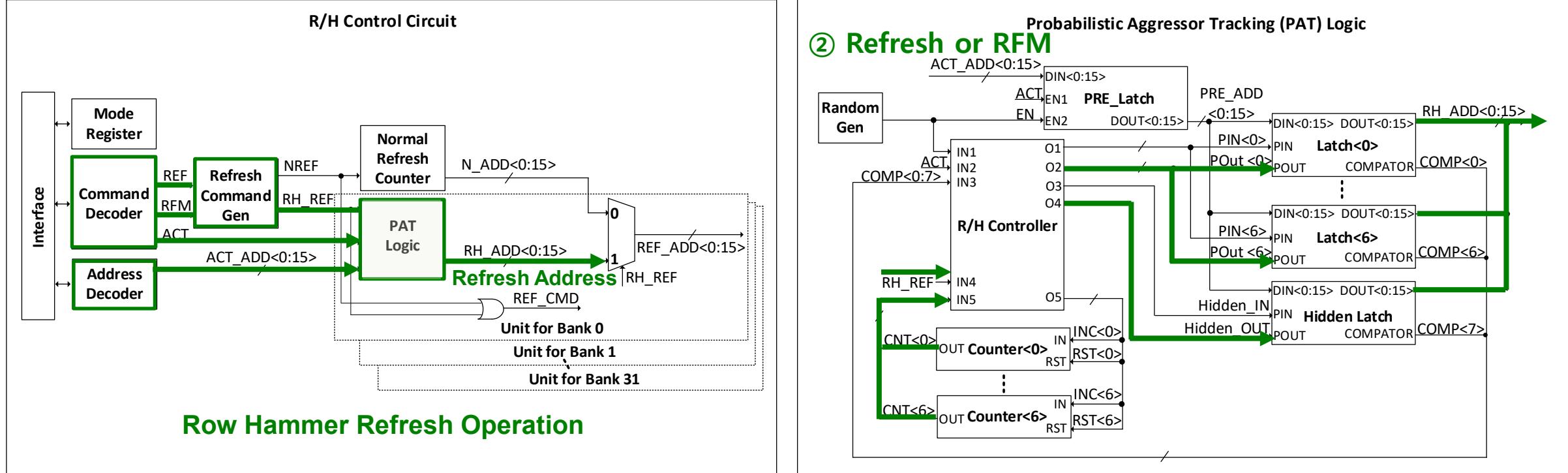
- Probabilistic approach to improve aggressor tracking accuracy
- Cost-effective since it is implemented within the peripheral area

Probabilistic Aggressor Tracking (PAT)



- Probabilistic approach to improve aggressor tracking accuracy
- Cost-effective since it is implemented within the peripheral area

Probabilistic Aggressor Tracking (PAT)

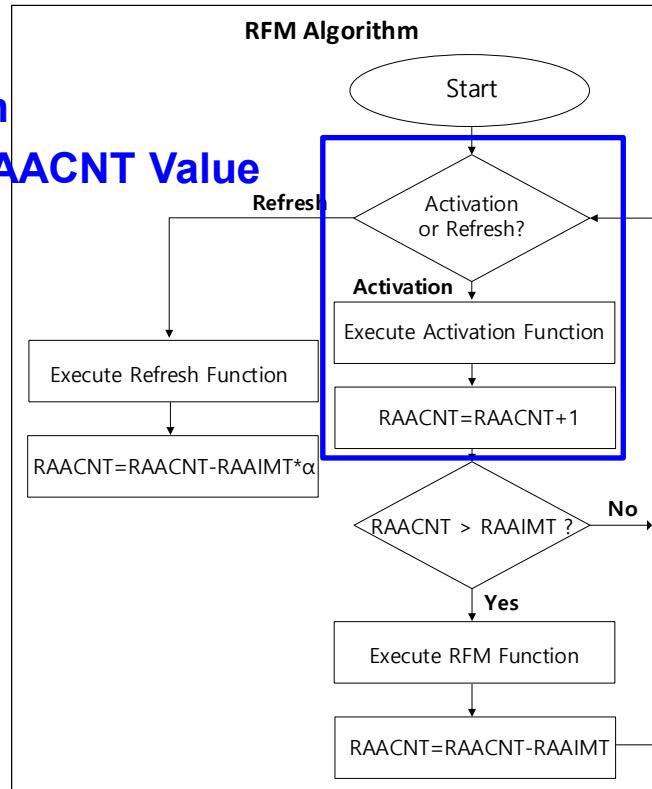


- Probabilistic approach to improve aggressor tracking accuracy
- Cost-effective since it is implemented within the peripheral area

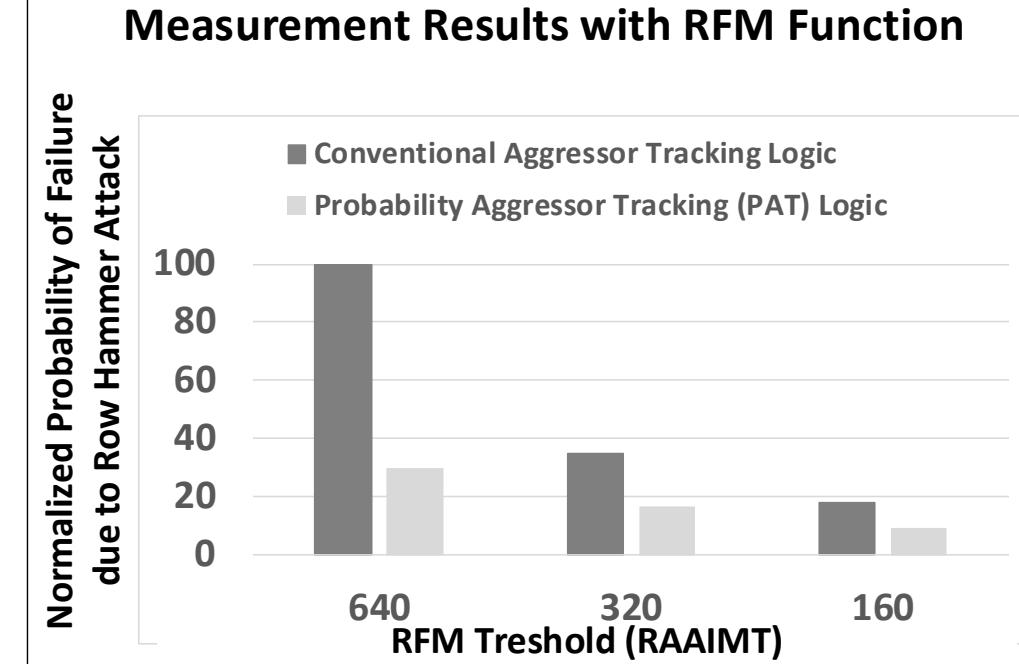
Refresh Management Function (RFM)

① Activation

- Increase RAACNT Value

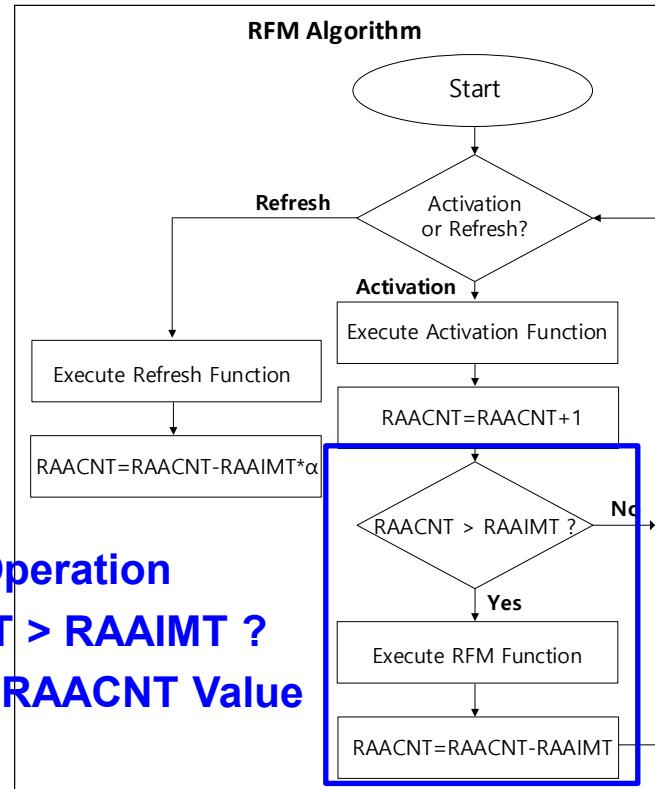


Measurement Results with RFM Function

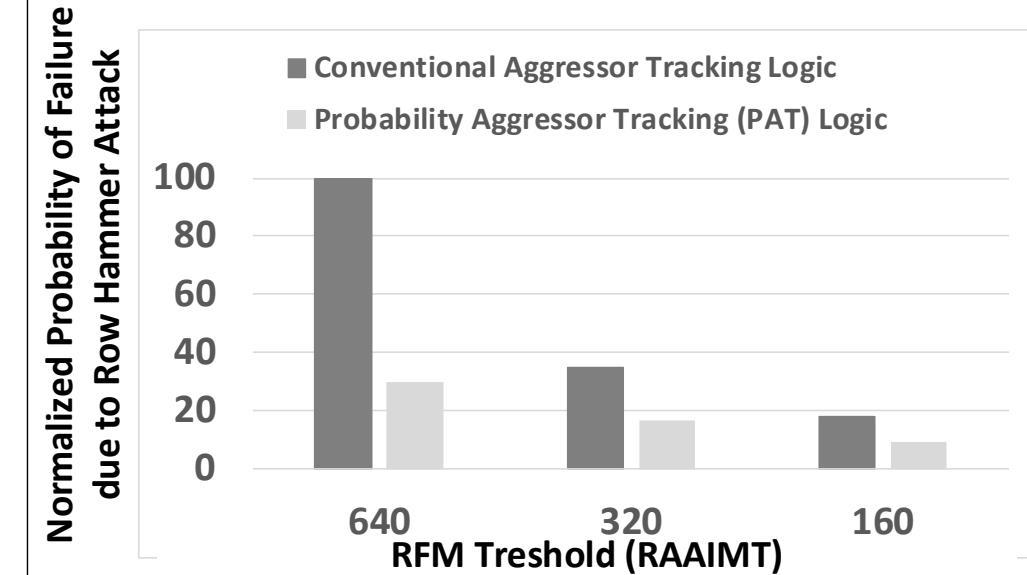


■ Execute RFM command when a malicious attack is detected

Refresh Management Function (RFM)



Measurement Results with RFM Function

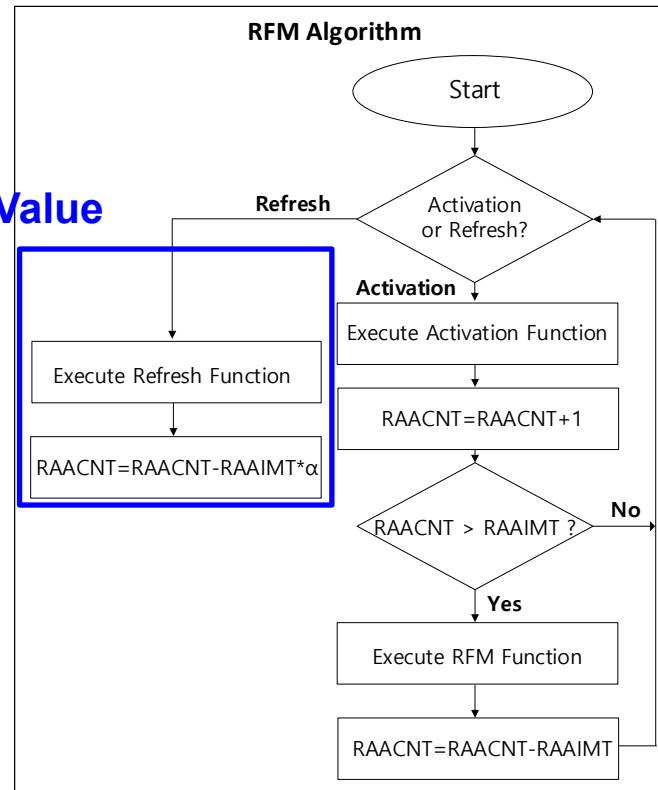


■ Execute RFM command when a malicious attack is detected

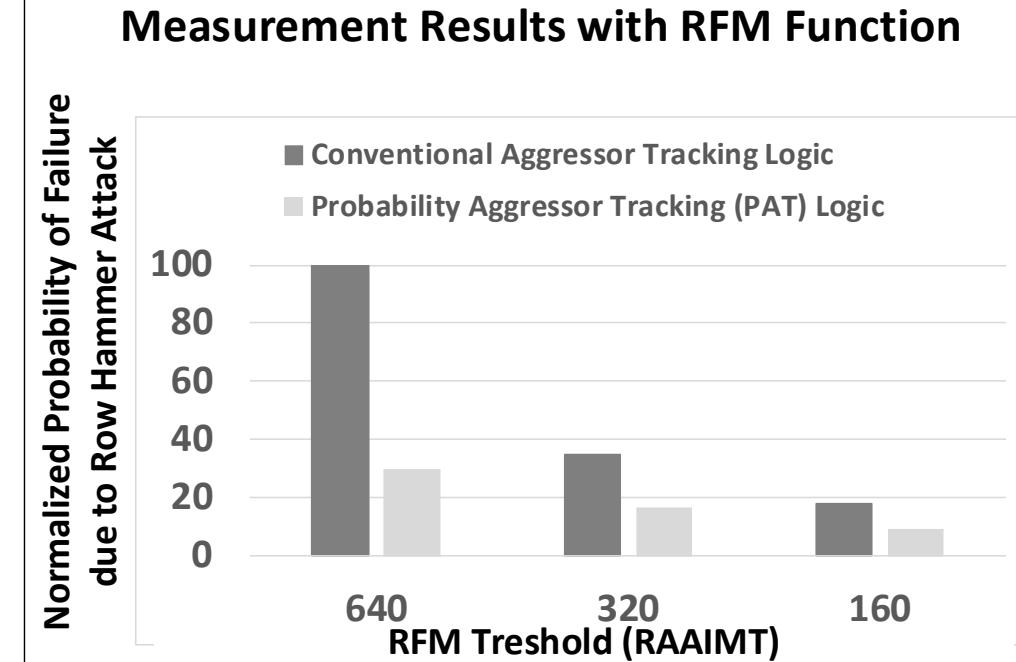
Refresh Management Function (RFM)

③ Normal Refresh

- Reduce RAACNT Value

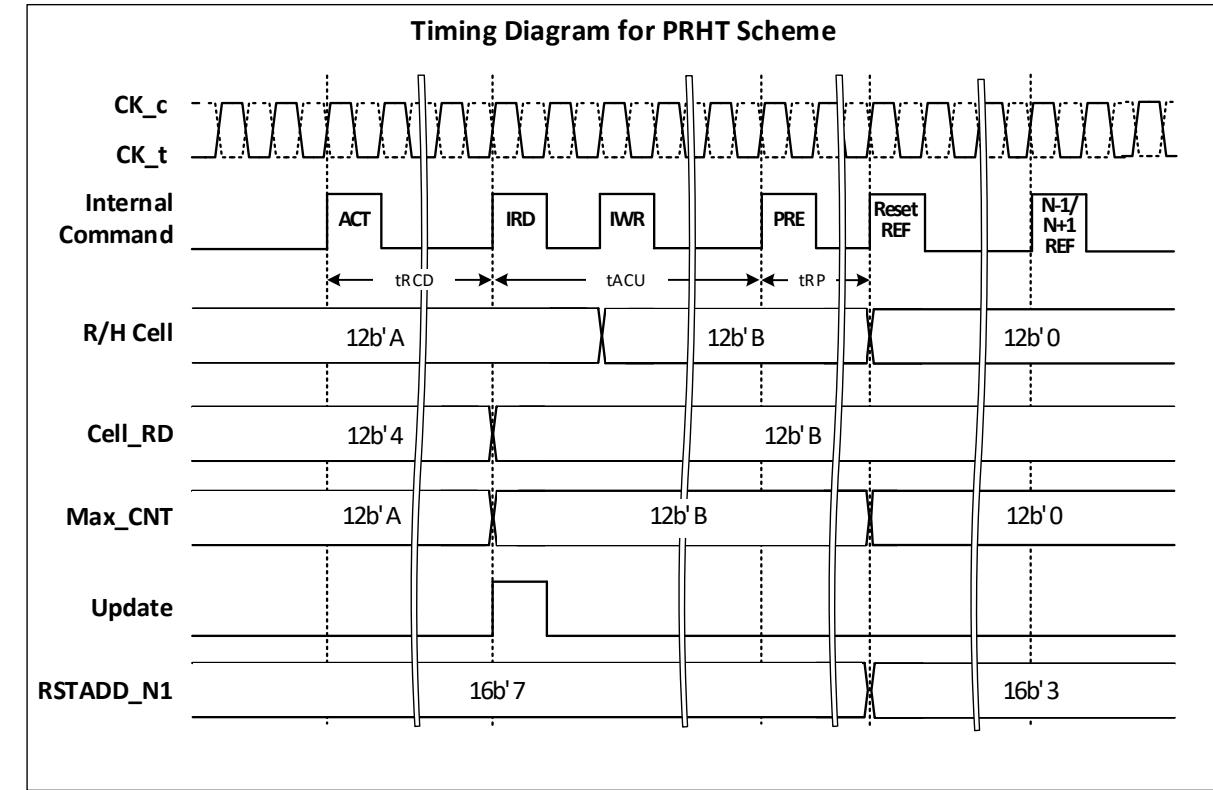
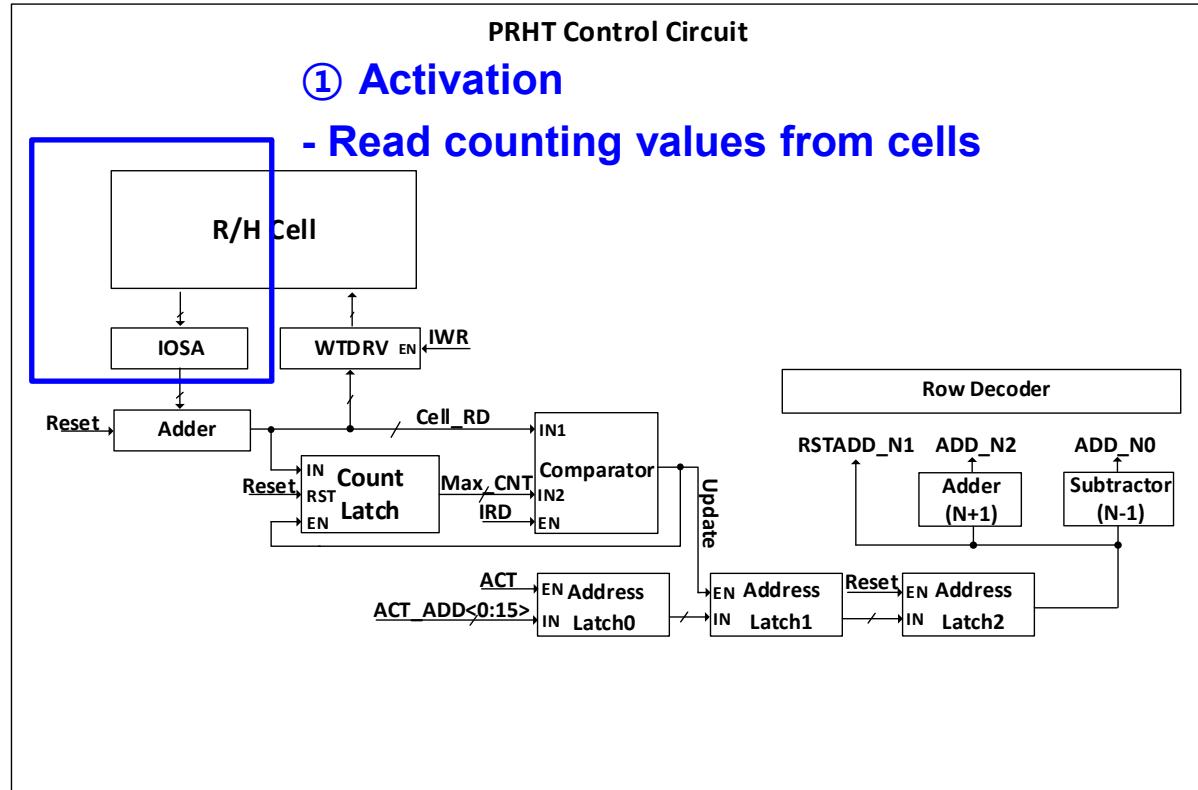


Measurement Results with RFM Function



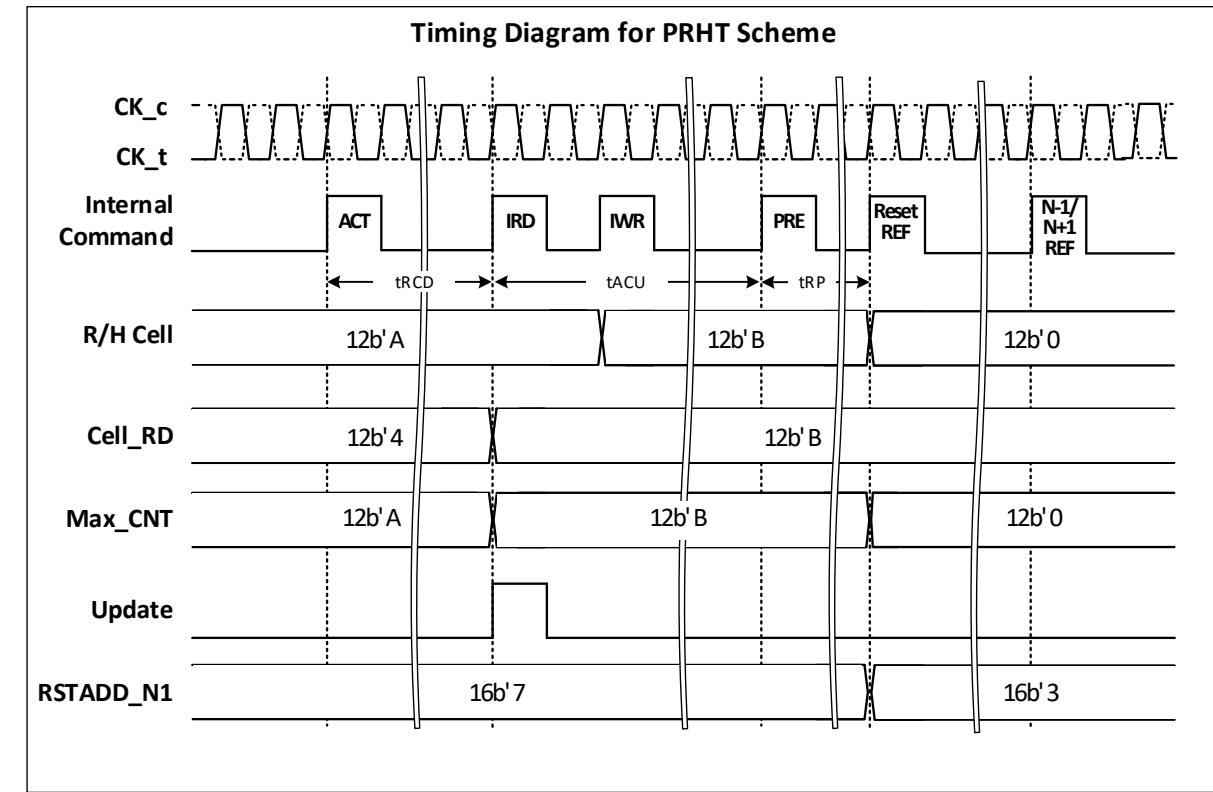
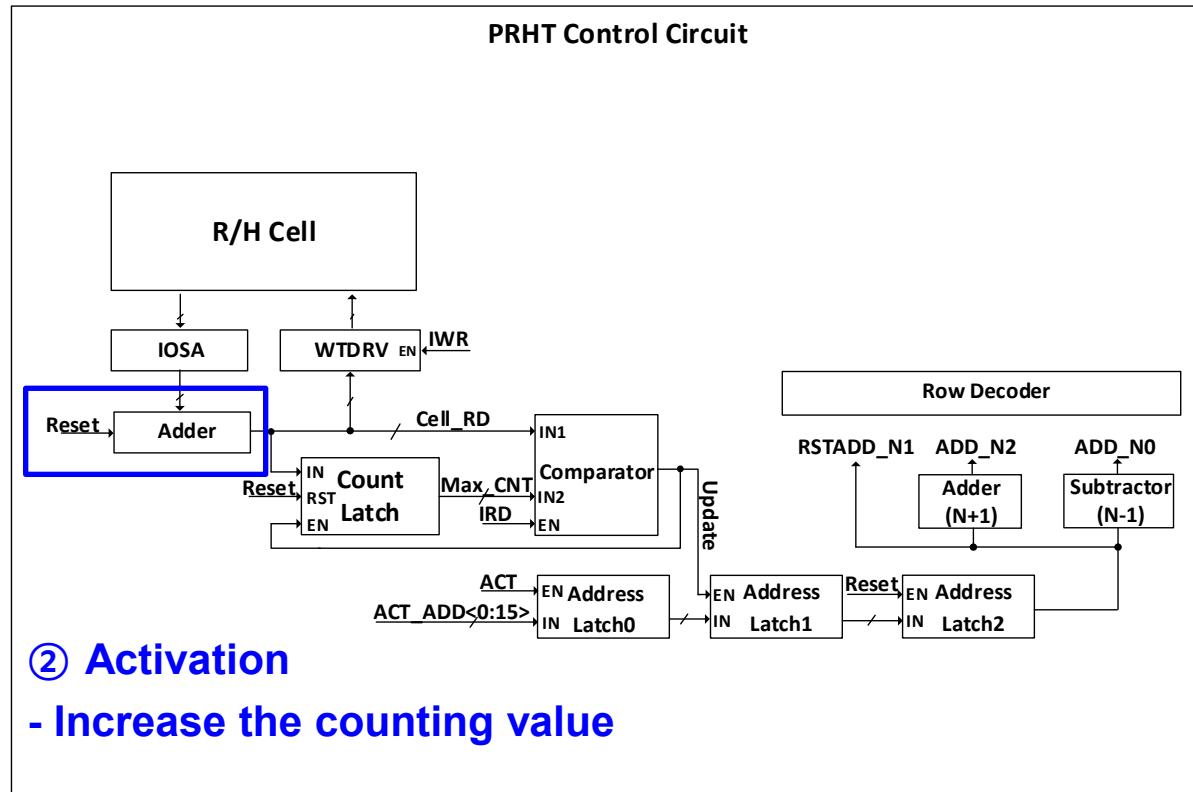
■ Execute RFM command when a malicious attack is detected

Per-Row Hammer Tracking (PRHT)



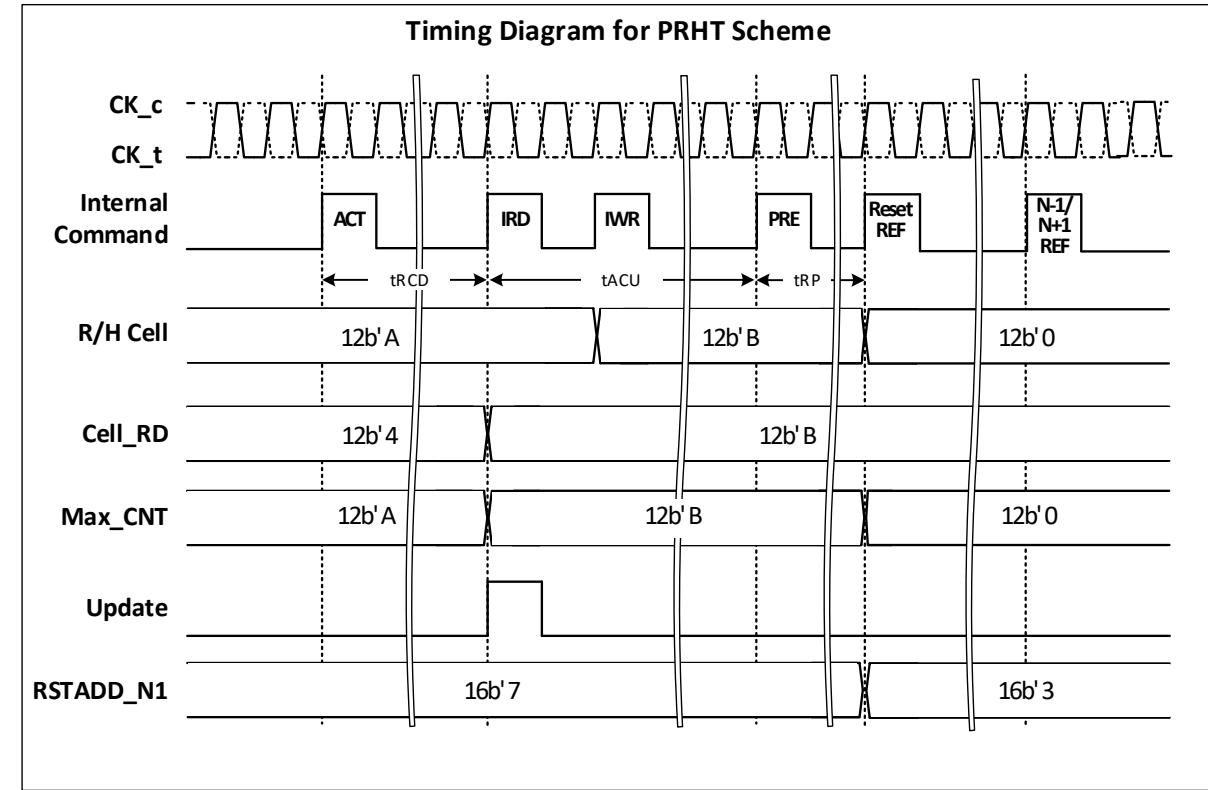
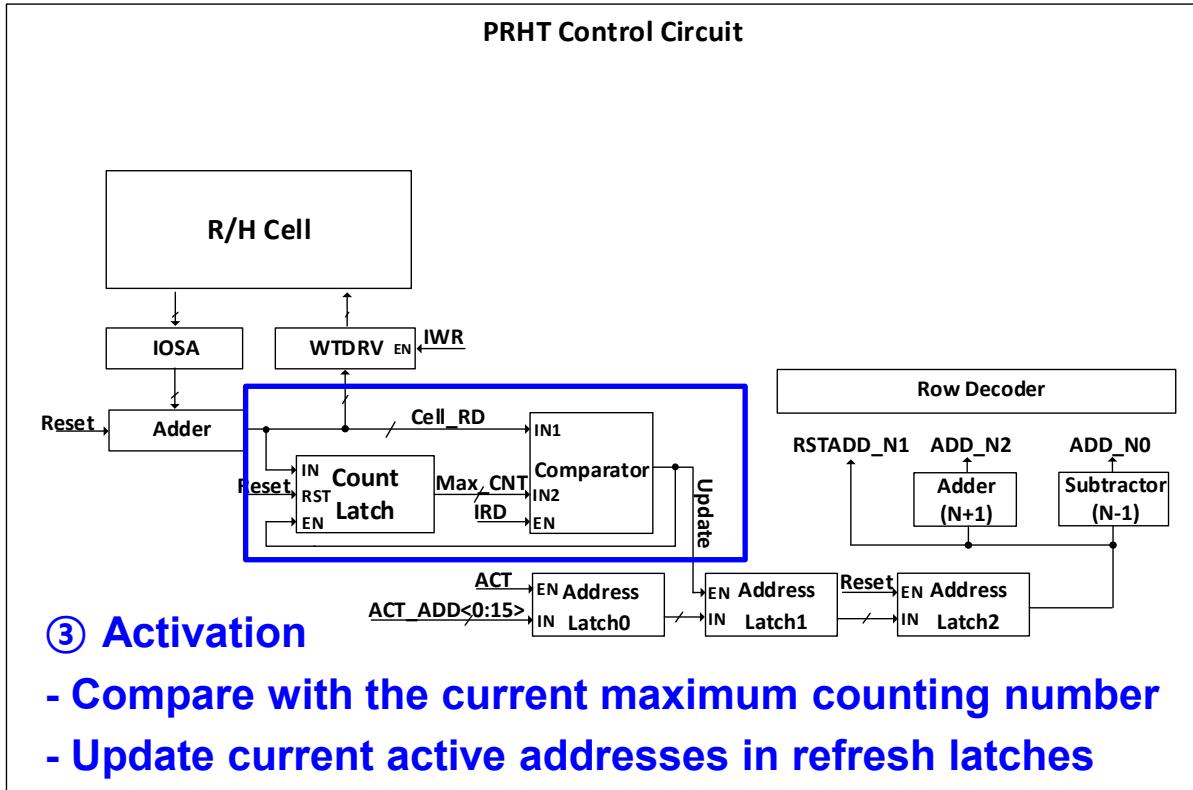
- Deterministic tracking scheme with additional bank area
- Count the number of active commands for each WL

Per-Row Hammer Tracking (PRHT)



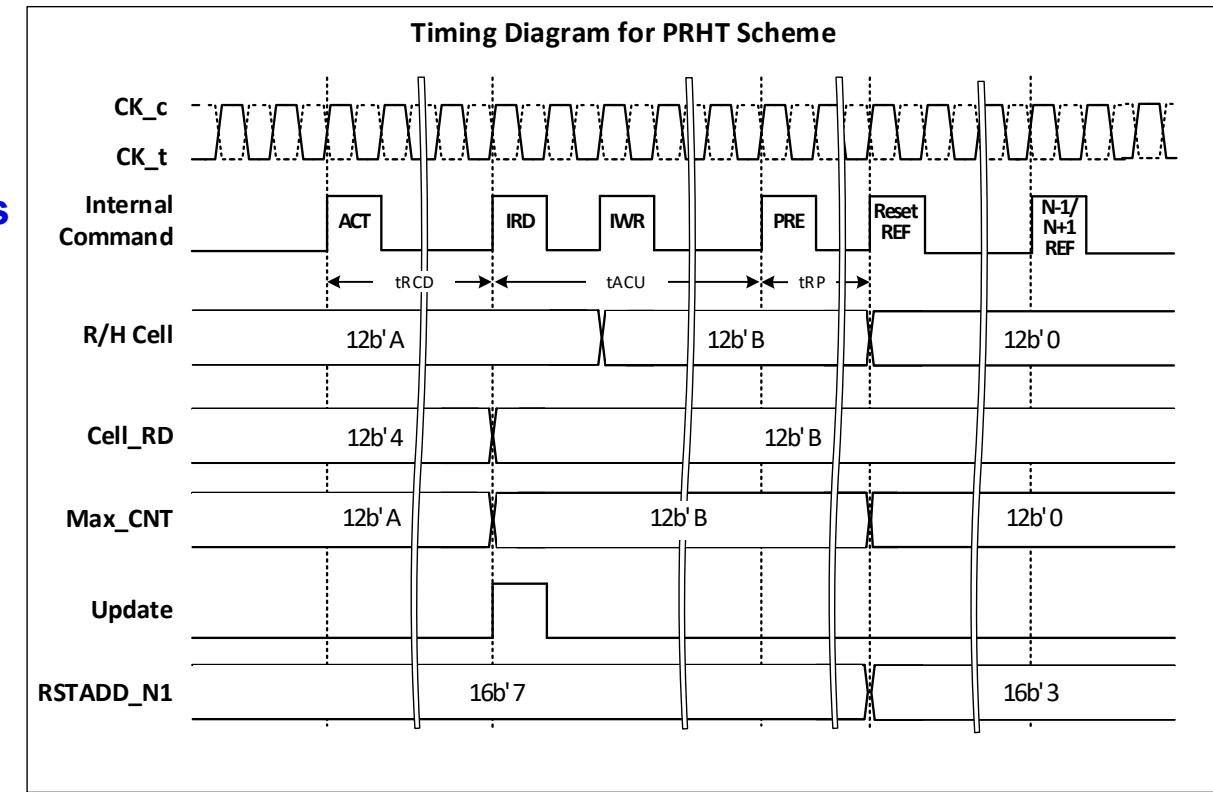
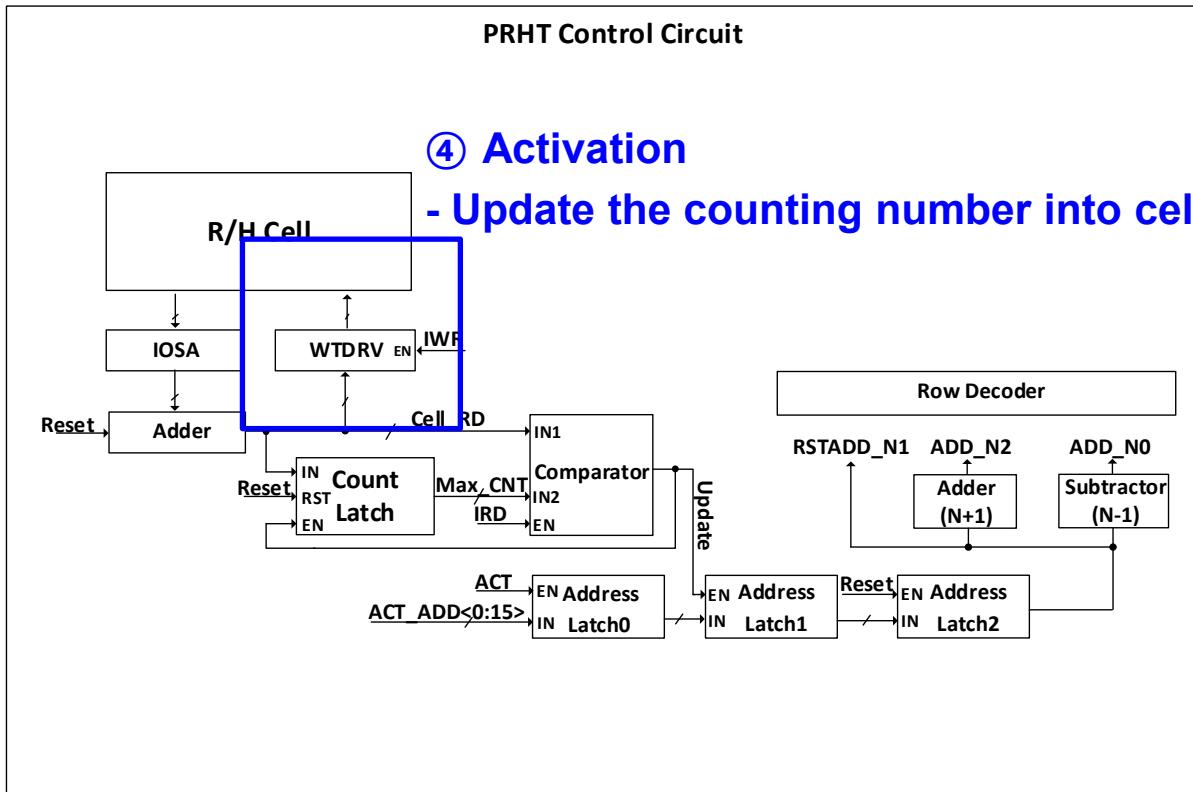
- Deterministic tracking scheme with additional bank area
- Count the number of active commands for each WL

Per-Row Hammer Tracking (PRHT)



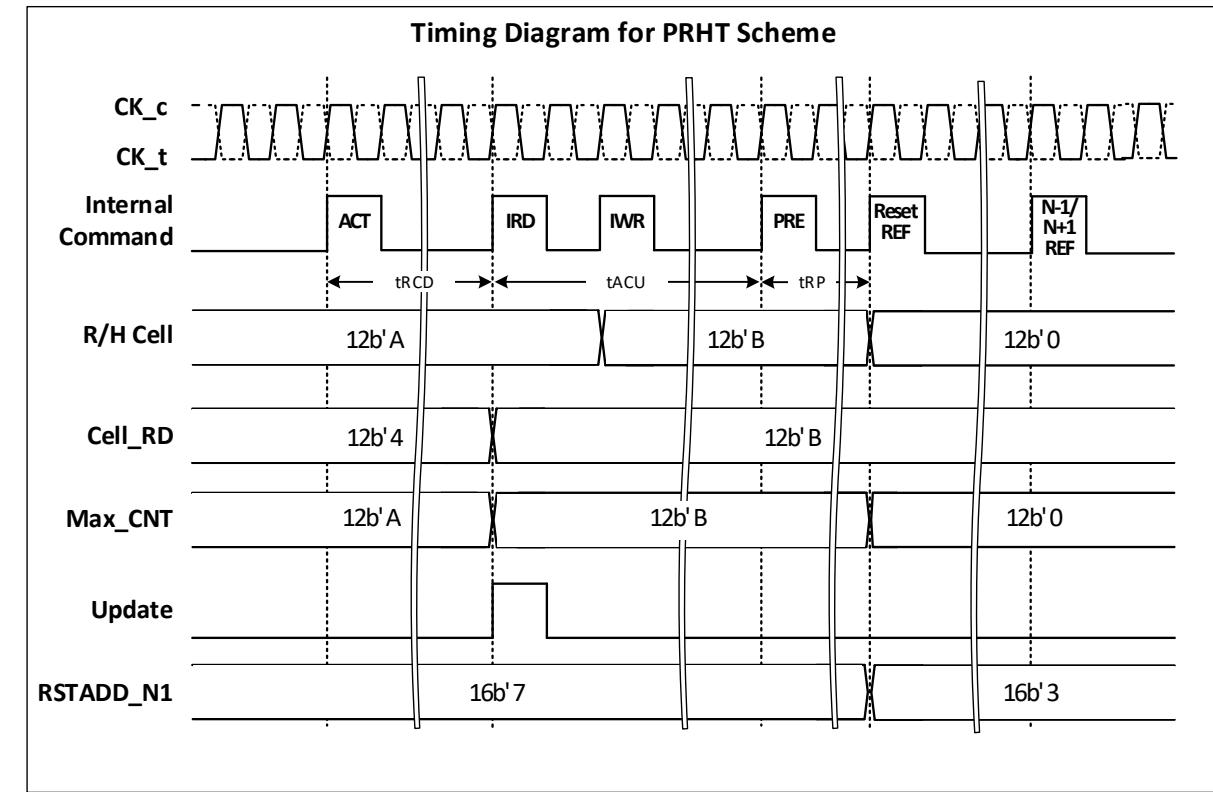
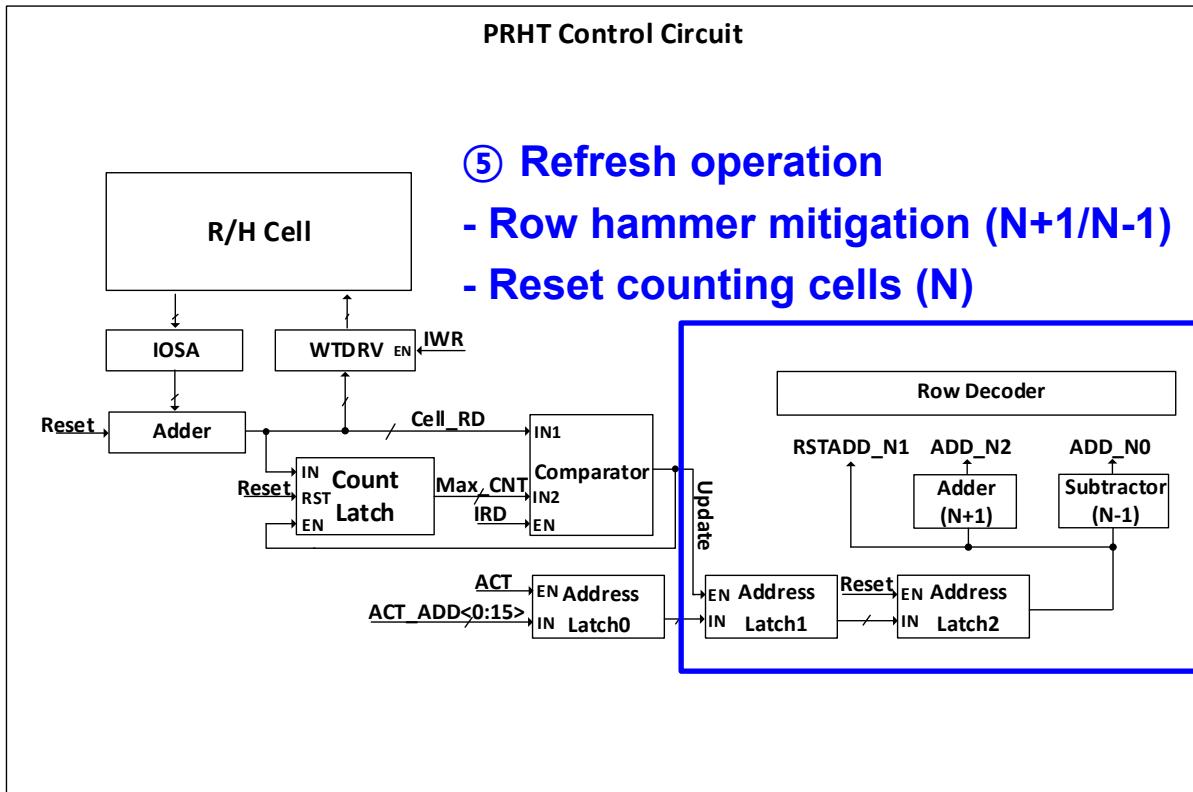
- Deterministic tracking scheme with additional bank area
- Count the number of active commands for each WL

Per-Row Hammer Tracking (PRHT)



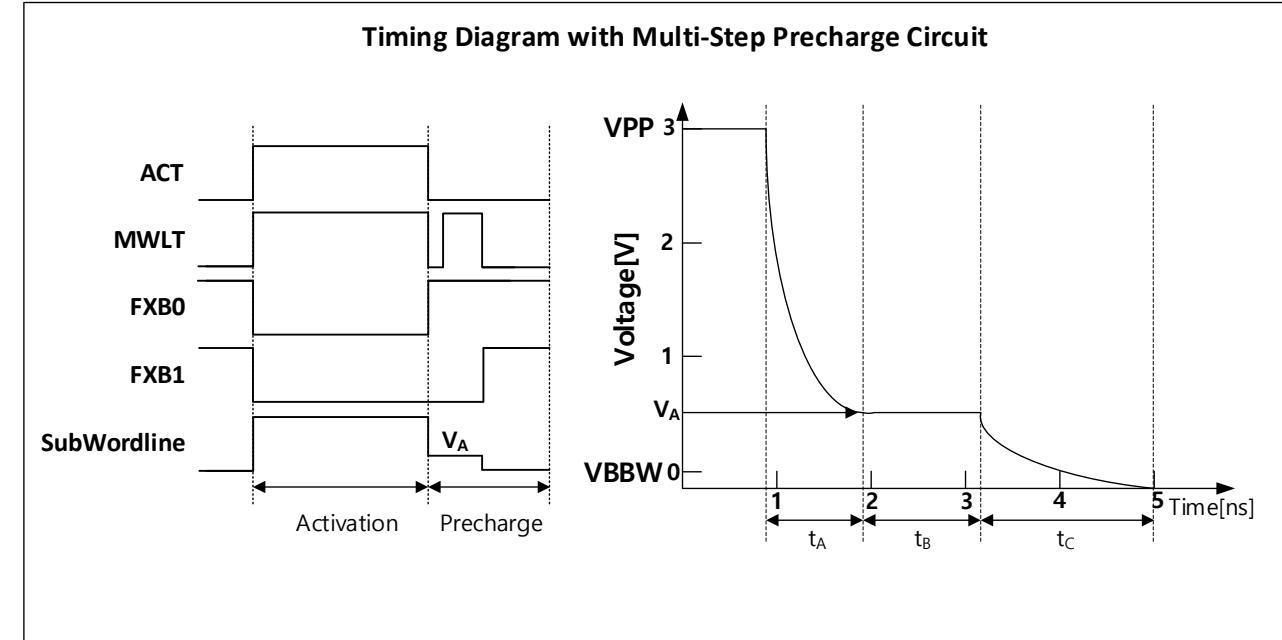
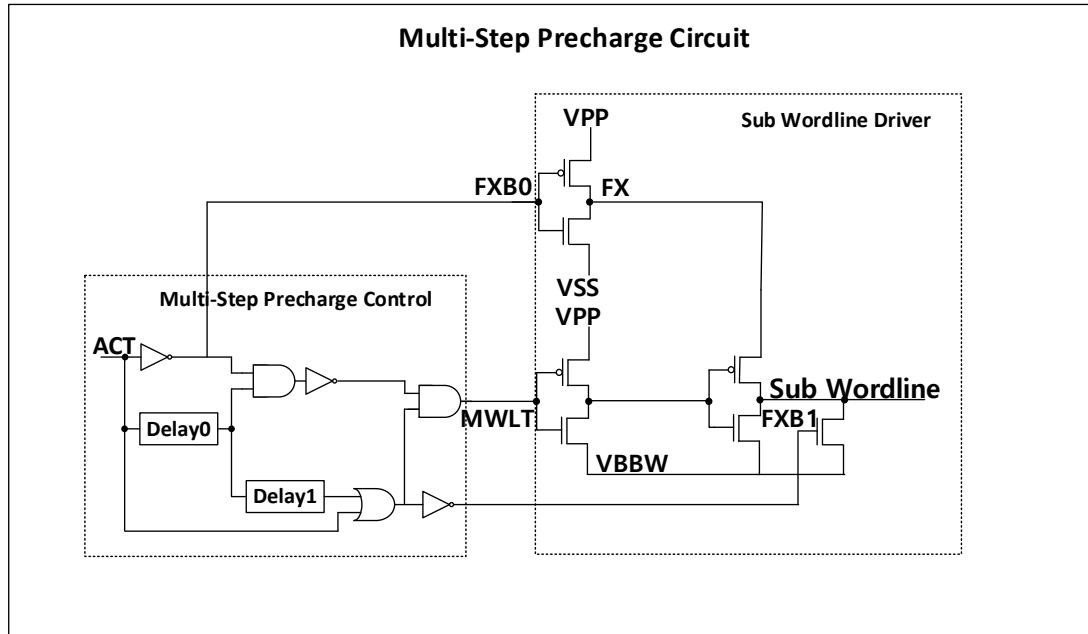
- Deterministic tracking scheme with additional bank area
- Count the number of active commands for each WL

Per-Row Hammer Tracking (PRHT)



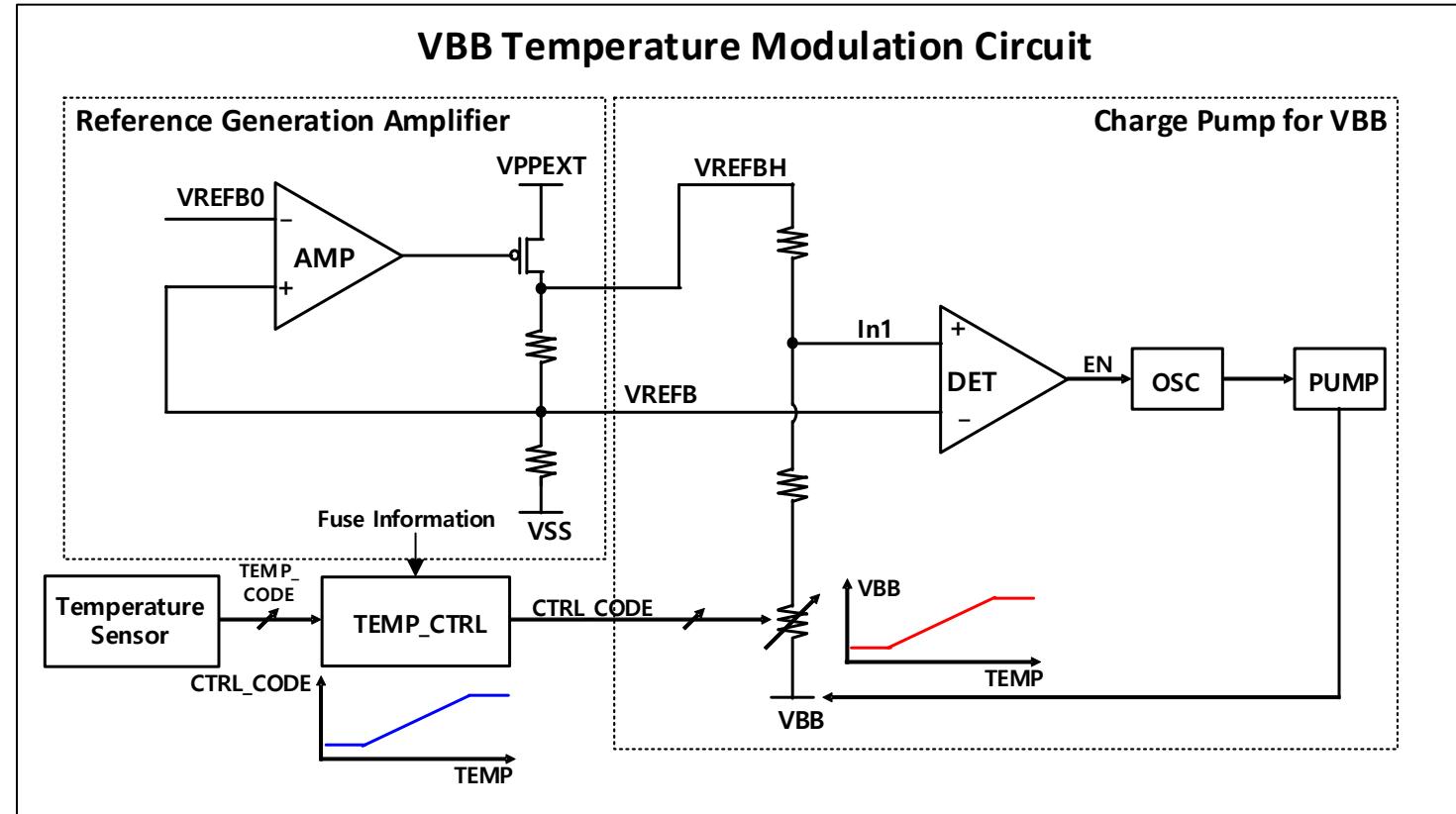
- Deterministic tracking scheme with additional bank area
- Count the number of active commands for each WL

Multi-Step Precharge Circuit



- Create sub-WL level to minimize charge loss for the victim cell
- The intrinsic row hammer tolerance is improved by 37%

Core-Bias Modulation

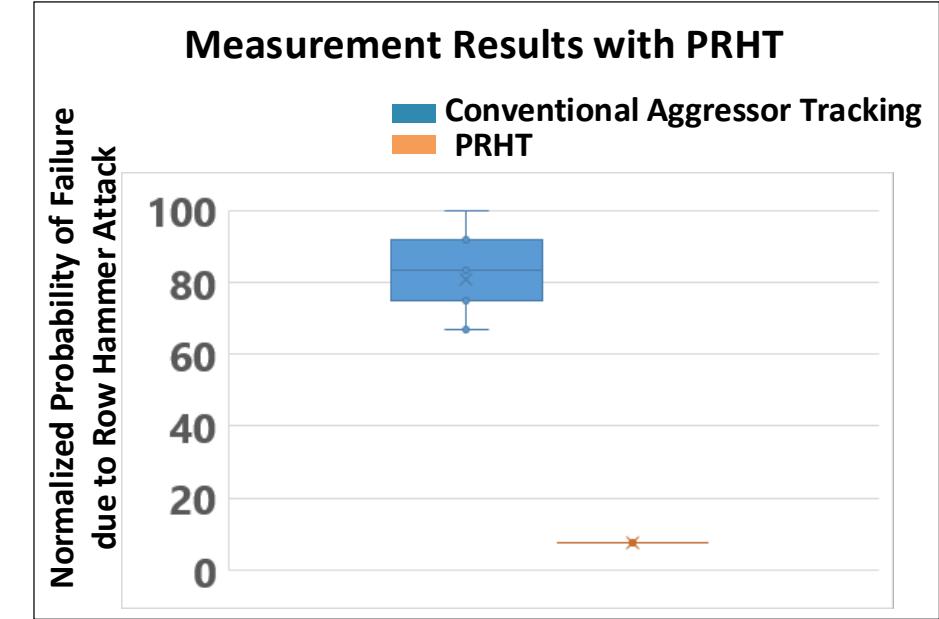
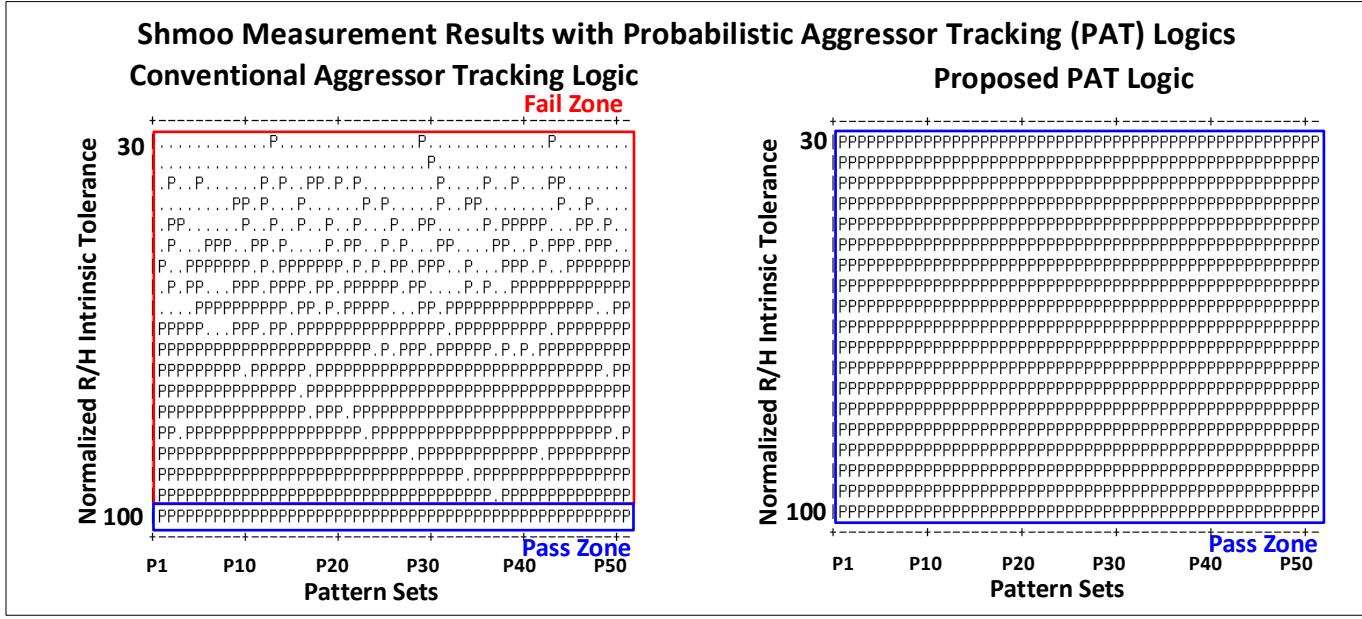


- VBB temperature-modulation circuit to maximize refresh retention time across 25 – 90°C by making the intrinsic row-hammer tolerance similar across temperature

Outline

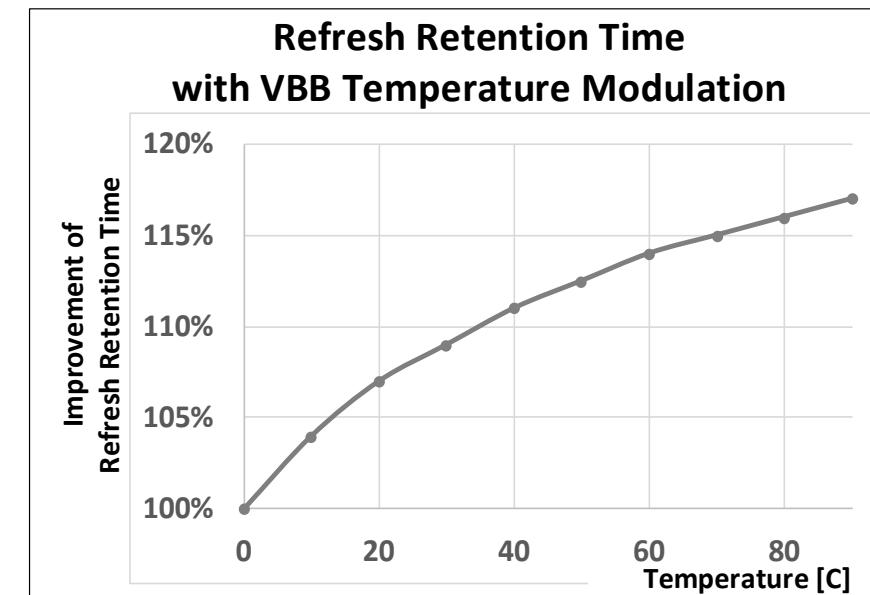
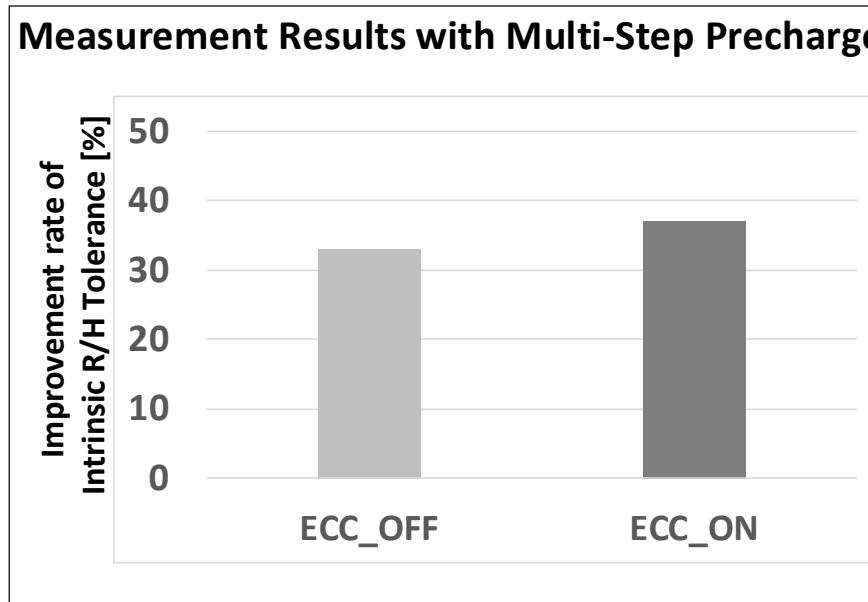
- **Introduction**
- **Architecture Overview**
- **Key Schemes**
 - Probabilistic aggressor tracking (PAT)
 - Refresh management function (RFM)
 - Per-row hammer tracking (PRHT)
 - Multi-step precharge circuit
 - Core-bias modulation
- **Measurement Results and Chip Implementation**
- **Conclusion**

Measurement Results



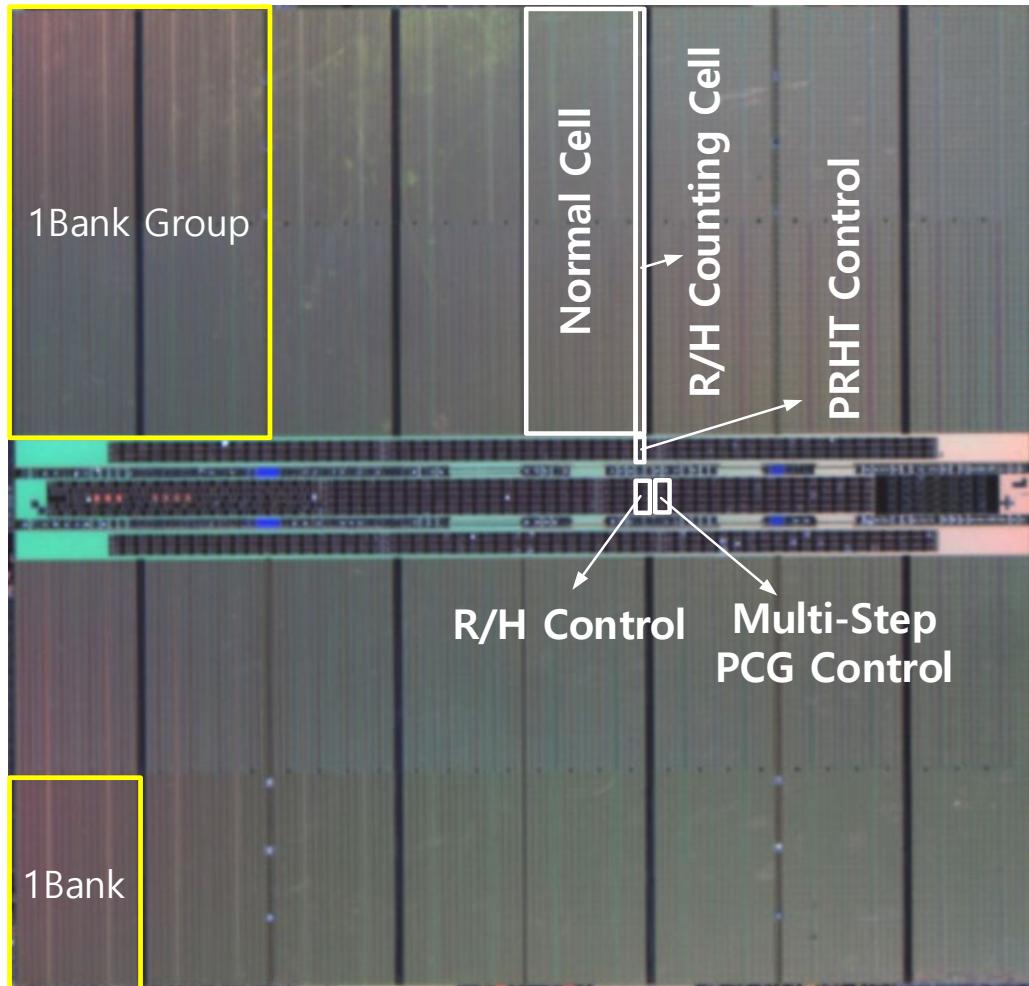
- PAT logic functionally passes row-hammer malicious pattern attacks even with a **66% lower intrinsic row-hammer tolerance**
- PRHT leads to a reduced probability of failure by **93.1%**

Measurement Results



- The intrinsic row hammer tolerance is improved by **37%** with the multi-step precharge scheme
- VBB modulation improves refresh retention time by **17%** at 90°C

Chip Implementation



Technology	1nm 5-metal DRAM HKMG process
Data Rate	6.4 Gbps/pin
Burst Length	BC8, BL16 on the fly
Number of IO	X4/X8/X16
Chip Size	7.159mm X 7.44mm
Supply Voltage	VDD/VDDQ 1.1V, VPP 1.8V
RAS Feature	In-DRAM ECC, PRHT

Conclusion

- A 16Gb DDR5 DRAM with row hammer protection and refresh management schemes is introduced
- 5 key design schemes are introduced
 - Probabilistic aggressor tracking (PAT)
 - Refresh management function (RFM)
 - Per-row hammer tracking (PRHT)
 - Multi-step precharge circuit
 - Core-bias modulation
- The proposed scheme can be the key enabling technology to extend DRAM technologies from 1a-nm to sub 10-nm