

A 16nm 256-bit Wide 89.6GByte/s Total Bandwidth In-Package Interconnect with 0.3V Swing and 0.062pJ/bit Power in InFO Package

Mu-Shan Lin, Chien-Chun Tsai, Cheng-Hsiang Hsieh,
Wen-Hung Huang, Yu-Chi Chen, Shu-Chun Yang, Chin-
Ming Fu, Hao-Jie Zhan, Jinn-Yeh Chien, Shao-Yu Li, Y.-H.
Chen, C.-C. Kuo, Shih-Peng Tai and Kazuyoshi Yamada

**Taiwan Semiconductor
Manufacturing Company, Ltd.**
DTP
Taiwan



Outline

- **Motivation**
 - In-Package Interconnect Applications
 - Advance Package Solutions
- **Introduction**
 - InFO Process
 - Low Power Design Concept (2013-VLSI)
- **System Architecture**
- **Circuit Descriptions**
- **Experimental Results**
- **Conclusion**

In-Package Interconnect Applications

- **LIPINCON™**: **L**ow-voltage-**I**n-**P**ackage-**I**Nter-**C**ONnect

- **High performance computing**

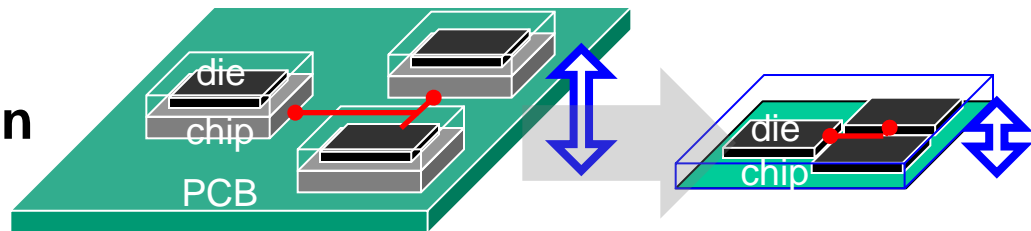
- ◆ Limit die size for yield

- **In package memory**

- ◆ Higher level cache

- **Heterogeneous integration**

- ◆ High speed SERDES

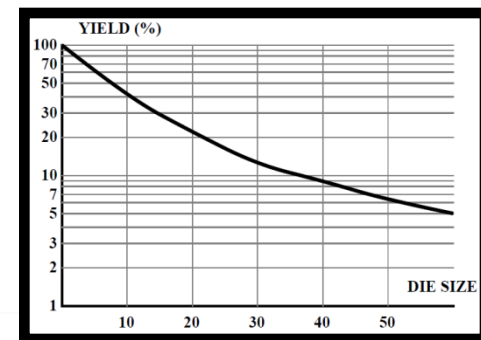
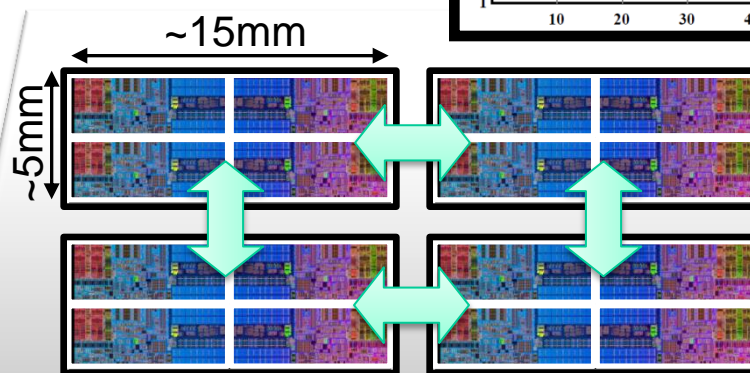
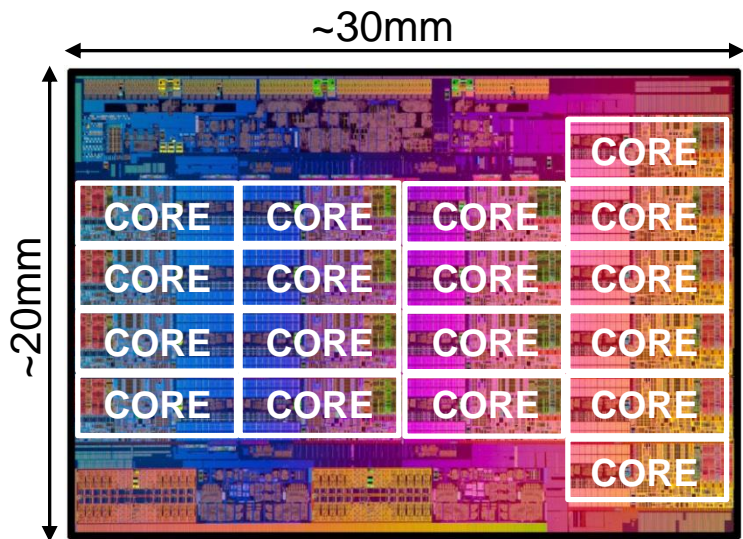


- ✓ **Smaller form factor**

- ✓ **Shorter interconnect trace**

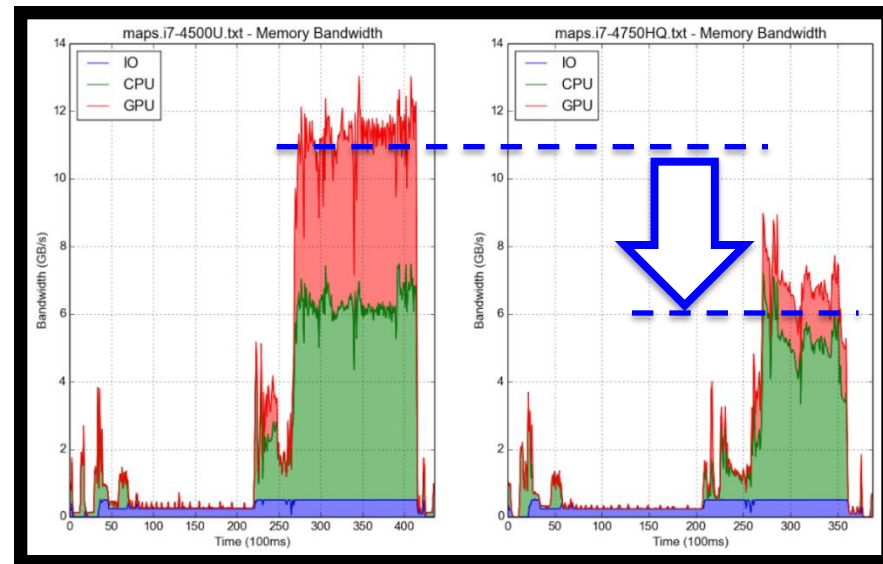
High Performance Computing (HPC)

- Yield goes down exponentially as die size gets large
 - Building large dice is quite difficult and very costly



In Package Memory (IPM)

- Additional memory hierarchy between on-chip SRAM and off-chip DDR
 - Smaller capacity; higher bandwidth; faster latency
- **Advantages**
 - DRAM is more cost-effective than eDRAM
 - Non-TSV allows fine pitch RDL and shorten trace between AP and IPM
 - Short trace enable termination-less IO design

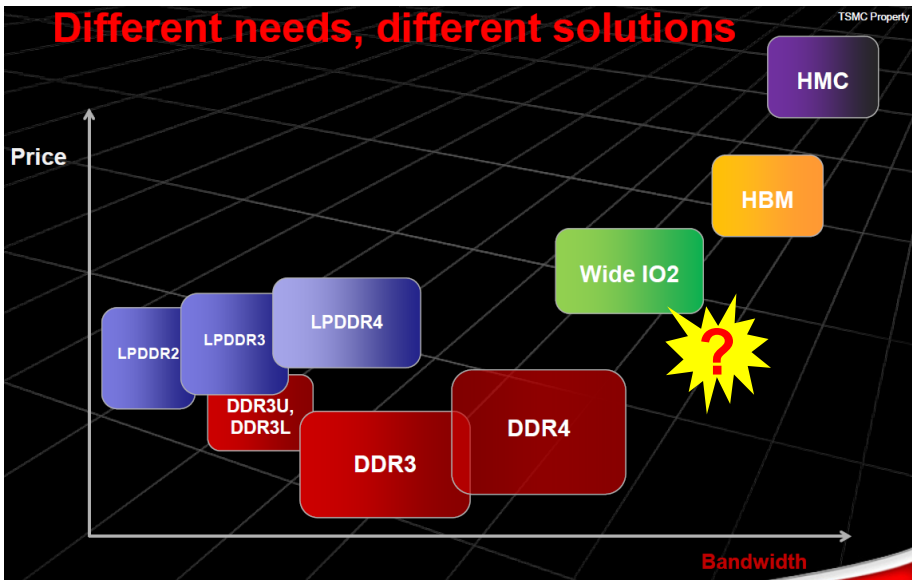
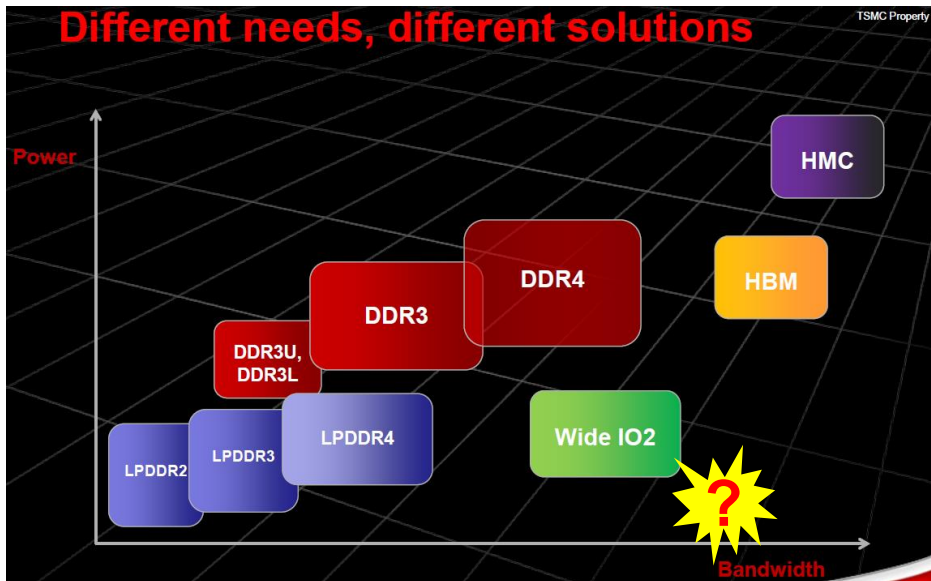


[2] Google, 2015-VLSI, “System Challenges..”

✓ **Save main DRAM bandwidth by 60%**

Memory Interface Development Trend

- Higher Bandwidth; Lower Power; Lower Cost (TSV-less)?



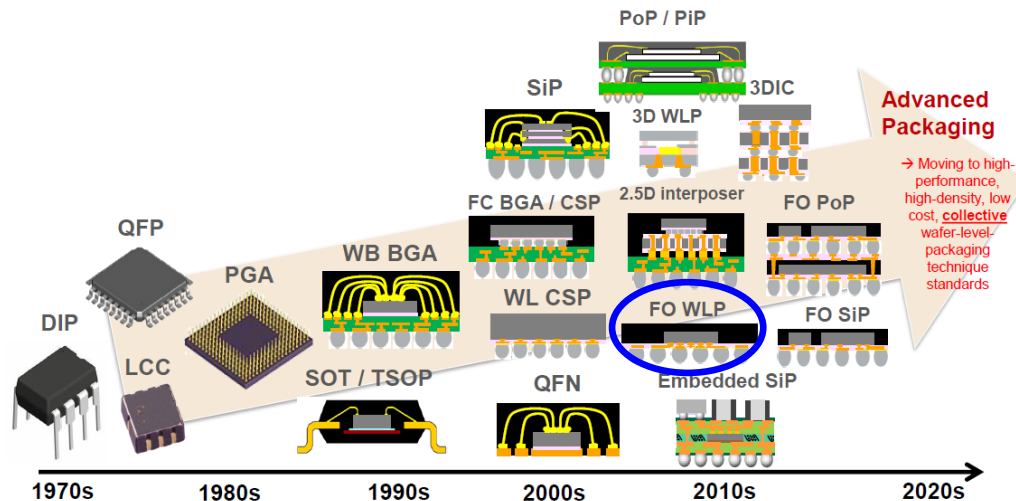
[3] TSMC, 2013-OIP “SoC Memory..”

Advanced Package Solutions

- Why we choose FO-WLP?
 - Ultra-fine pitch RDL (W/S < 5um/5um)
 - Ultra-thin package (~0.6mm including BGA)
 - Smaller die size (no-TSV)
 - Suit for smartphones, tablets, and wearables

Semiconductor chip packaging market evolution

- The pace of innovation in chip packaging industry has never been faster!
 - Today driven by semiconductor company giants (Intel, Samsung, TI, STMicro, TSMC, Qualcomm...) along with "Top 5" biggest packaging subcontractors (ASE, Amkor, SPIL, STATschippac, PTI...)



© 2012 • 2

Copyrights © Yole Développement SA. All rights reserved.

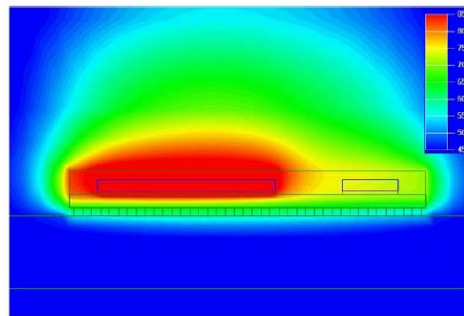
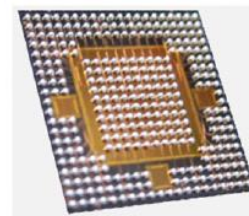
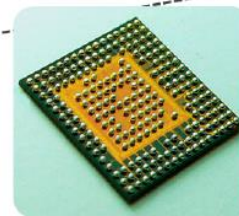
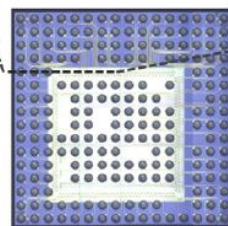
YOLE DÉVELOPPEMENT

[4] 2012 Business Update "3DIC & TSV interconnects"

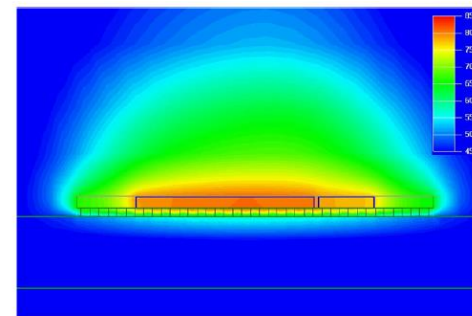
TSMC Package Solution

● InFO-WLP (INtegrated Fan-Out Wafer-Level-Package) vs FC-BGA

- More pin counts with smaller die size
- Lower parasitic resistance (thicker copper traces)
- Lower substrate loss (molding compound)
- Better thermal behavior (smaller form factor)



(a) Multi-Chip Module (MCM).



(b) InFO-WLP.

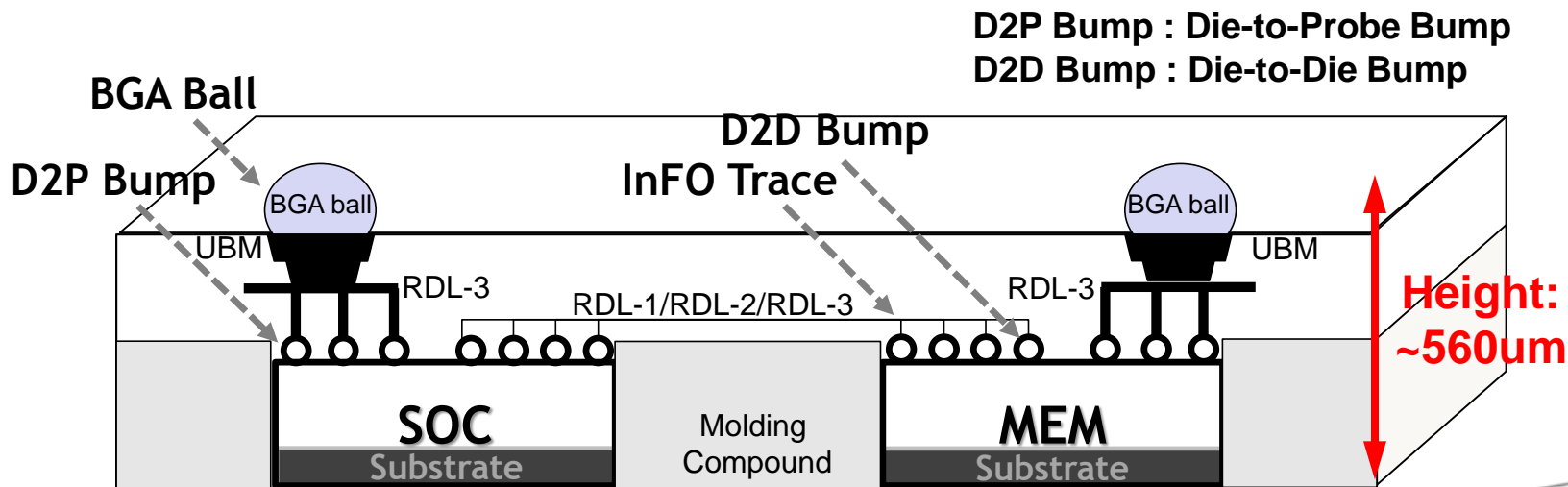
[5] TSMC, 2012-IEDM, "High-Performance.."

Project Target

- Demonstrate an in-package interconnect for in-package memory in InFO package
 - Less capable of memory process is considered
- Low-Power → Low-Latency → Small Area
 - Try to be transparent as possible
- 2Gbit/s/pin; 64Gbyte/s total bandwidth
- Power efficient IO
- Prompt and automatic timing-calibration scheme

Introduction

- TSMC InFO side-by-side
 - 560um thickness of packaged chip including 3*RDL & BGA

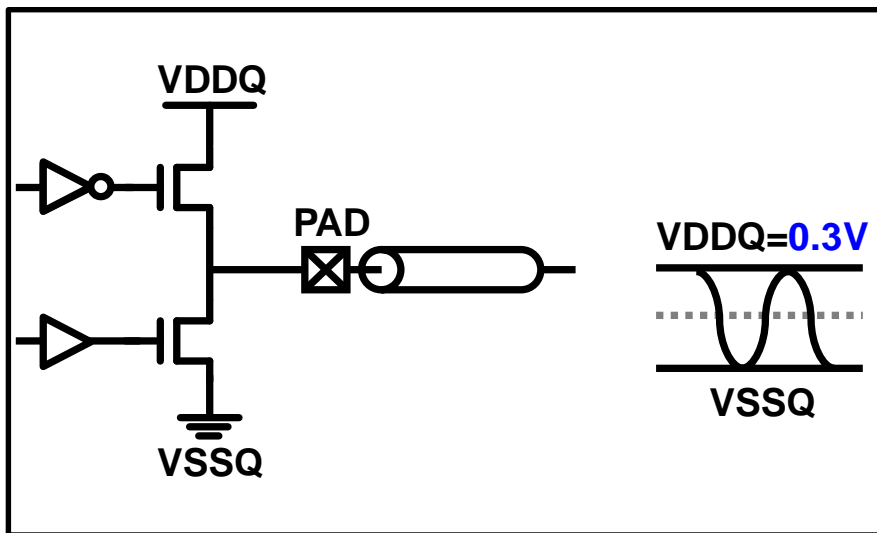


Introduction

- Low power design concept 2013-VLSI by TSMC

Reduced-power TX

- ❑ Low swing with lower VDDQ
- ❑ Termination-less
- ❑ Dynamic power only



[6] TSMC, 2013-VLSI, “An extra low-power..”

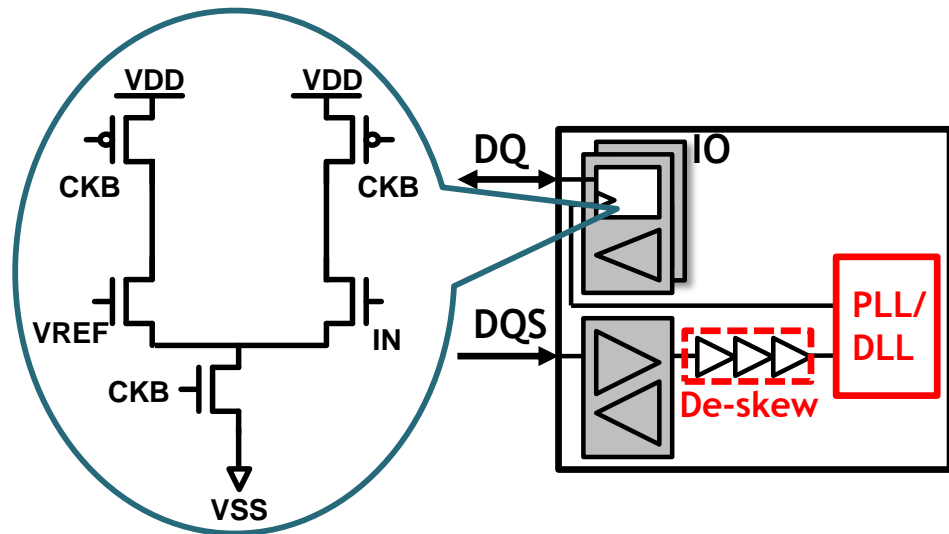
[7] JESD8-28

Introduction

- Low power design concept 2013-VLSI by TSMC

Reduced-power RX

- ❑ Clock-based sense amplifier
- ❑ No balance buffer
- ❑ Termination-less
- ❑ Dynamic power only



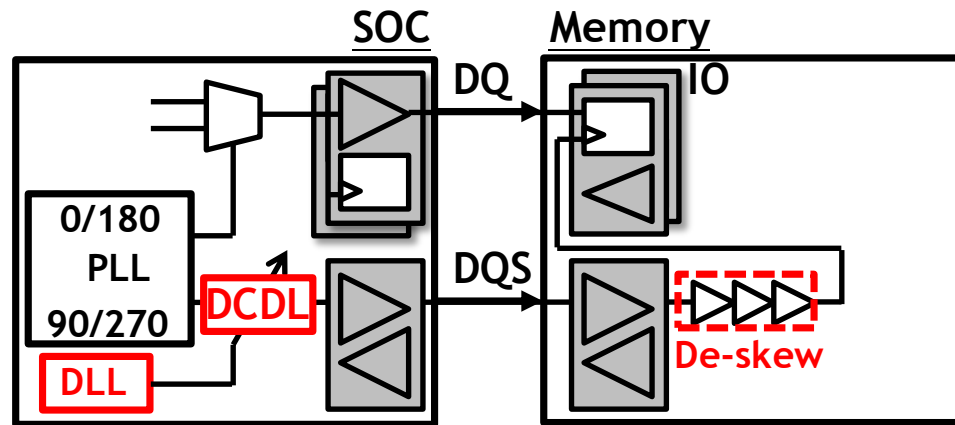
[6] TSMC, 2013-VLSI, "An extra low-power.."

Introduction

- Low power design concept 2013-VLSI by TSMC

Simplified design in DRAM

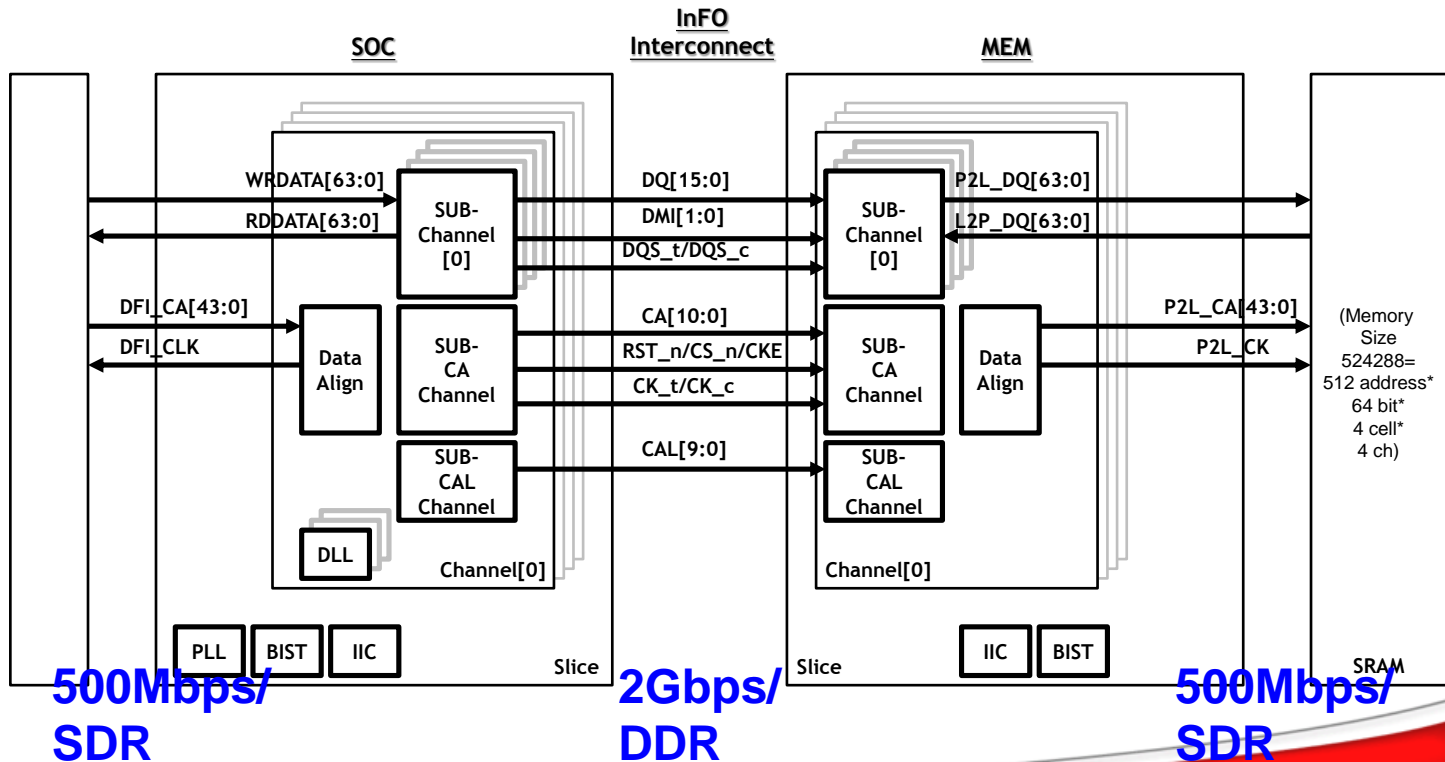
- Get rid of DLL/PLL in DRAM



[6] TSMC, 2013-VLSI, "An extra low-power.."

System Architecture

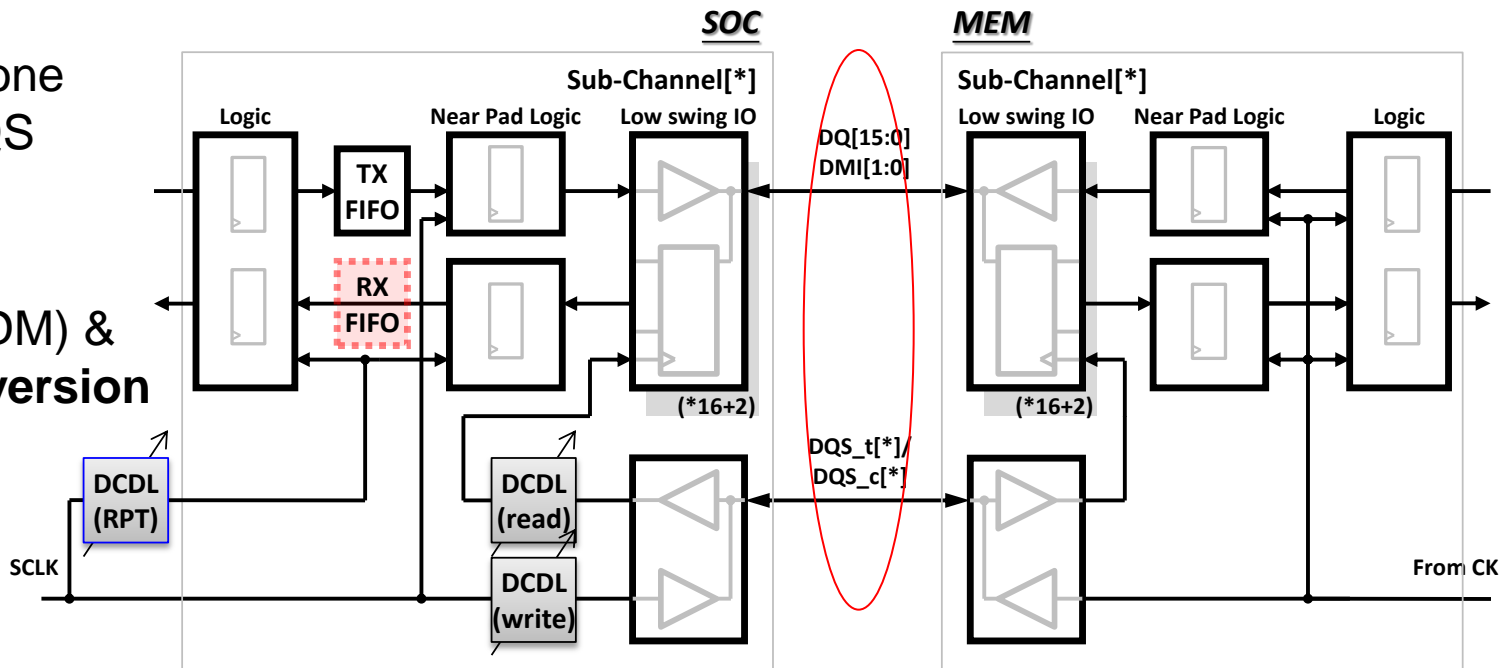
- ❑ 64GByte/s
- ❑ 256-DQ
- ❑ Bi-directional DQ/DQS
- ❑ Modular design
- ❑ Easily scalable
- ❑ Physical balance



Circuit Description

● Sub-DQ-channel architecture

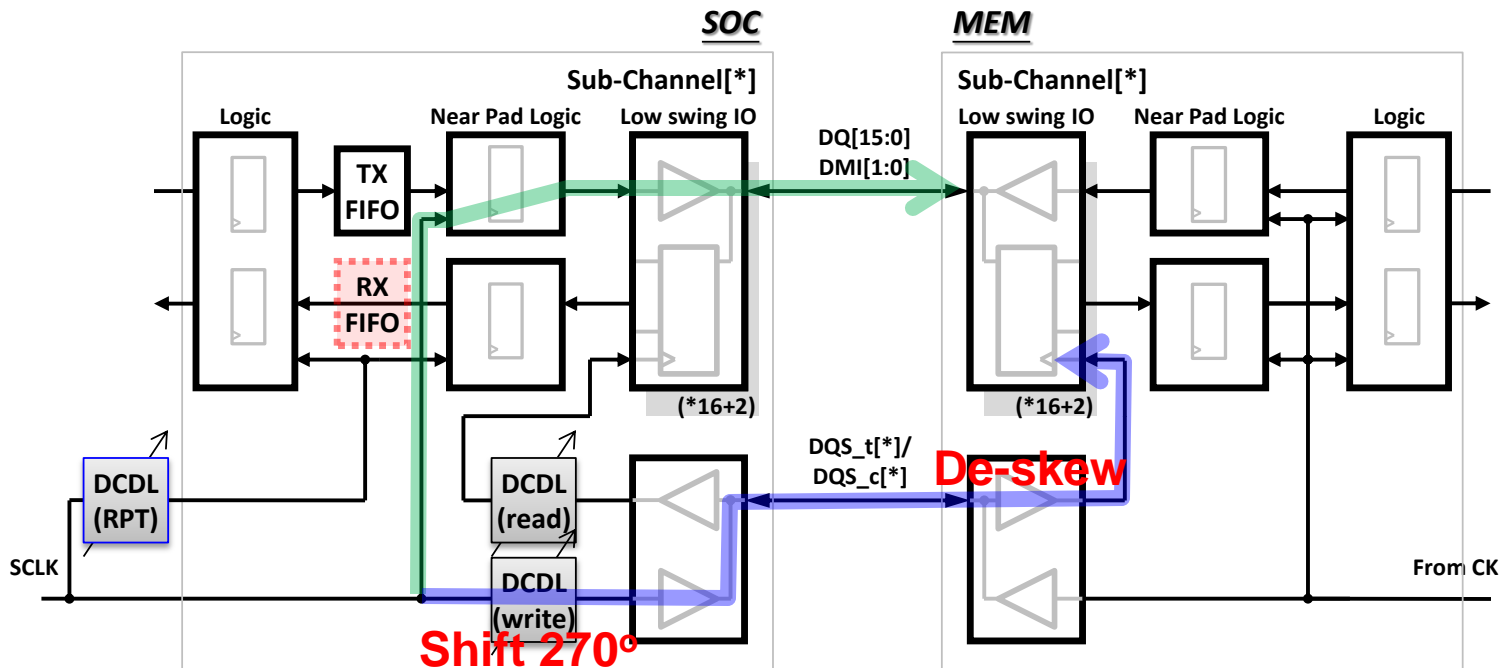
- ❑ 16-DQ share one differential DQS
- ❑ Per byte DMI combines **Data_mask (DM)** & **Data_bus_inversion (DBI)**



Circuit Description

- Sub-DQ-channel architecture

- WRITE-path
 - DLL-TX [6]



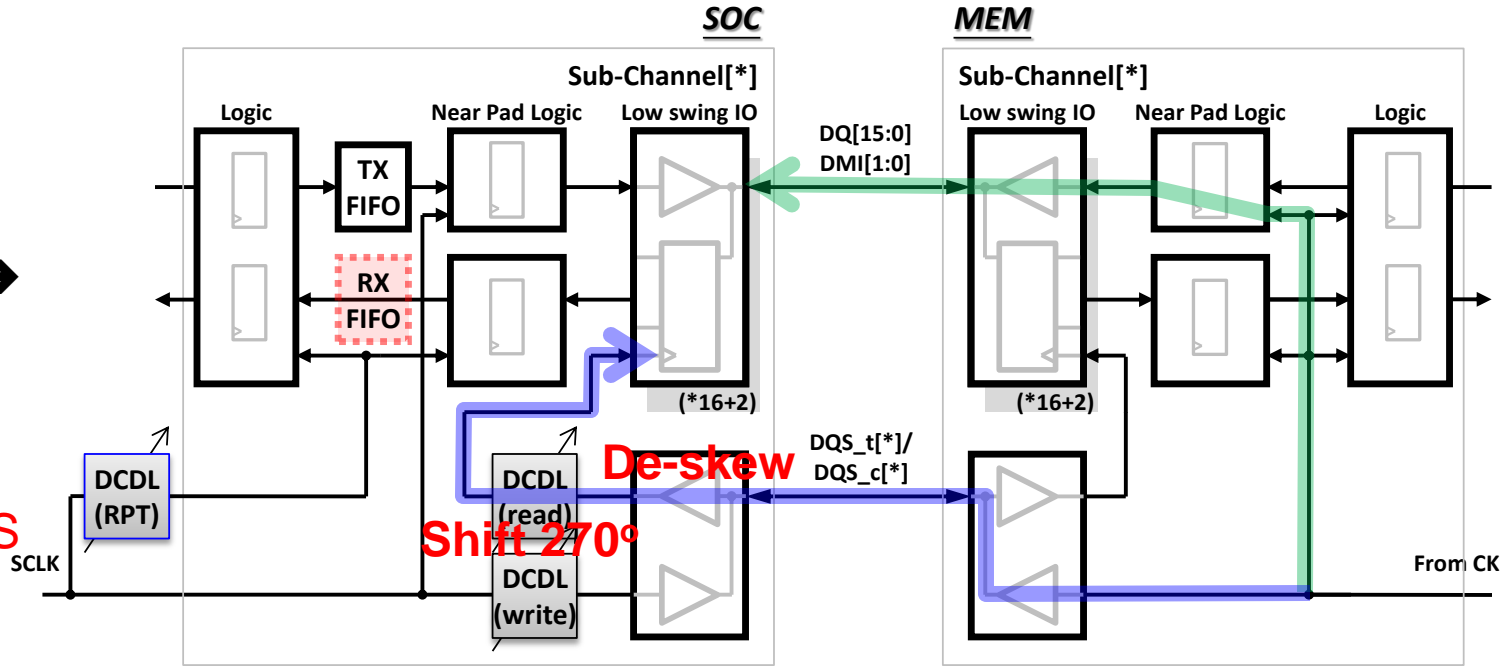
[6] TSMC, 2013-VLSI, "An extra low-power.."

Circuit Description

Sub-DQ-channel architecture

- READ-path
 - DLL-RX [6]
 - ~~RXFIFO~~ → DLL-RPT

- Asynchronous clock between SCLK & RDQS

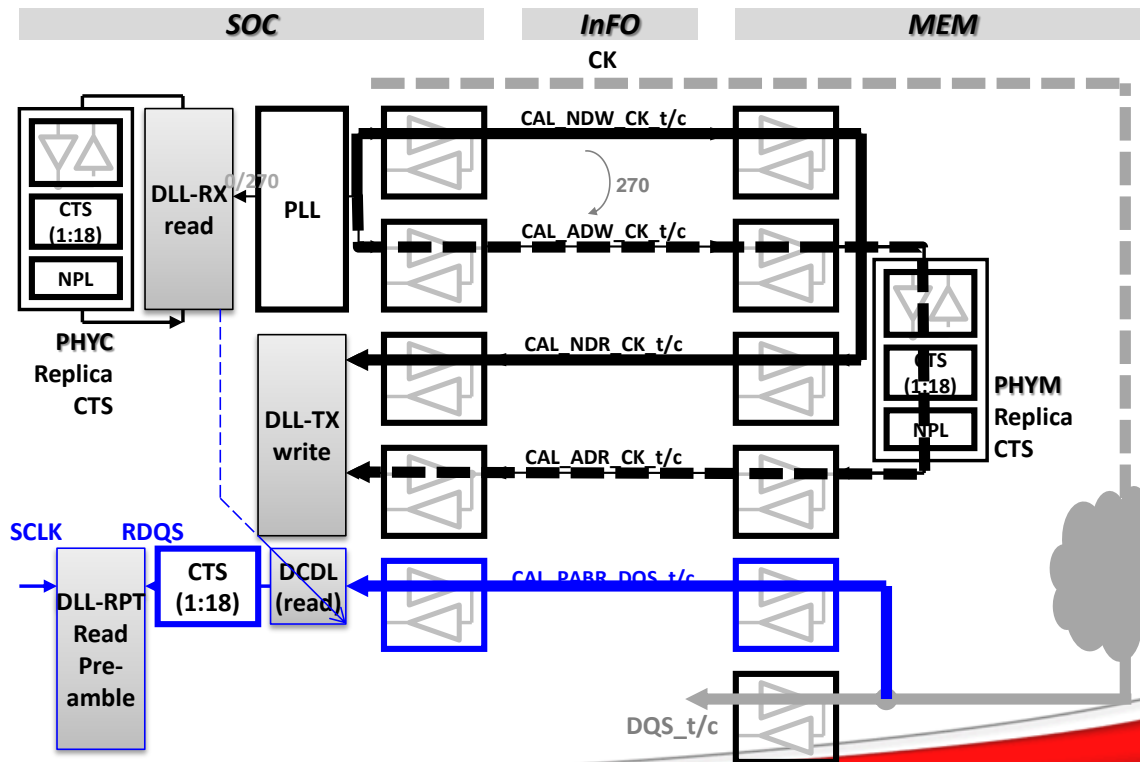


[6] TSMC, 2013-VLSI, "An extra low-power.."

Circuit Description

● Sub-CAL(calibration)-channel architecture

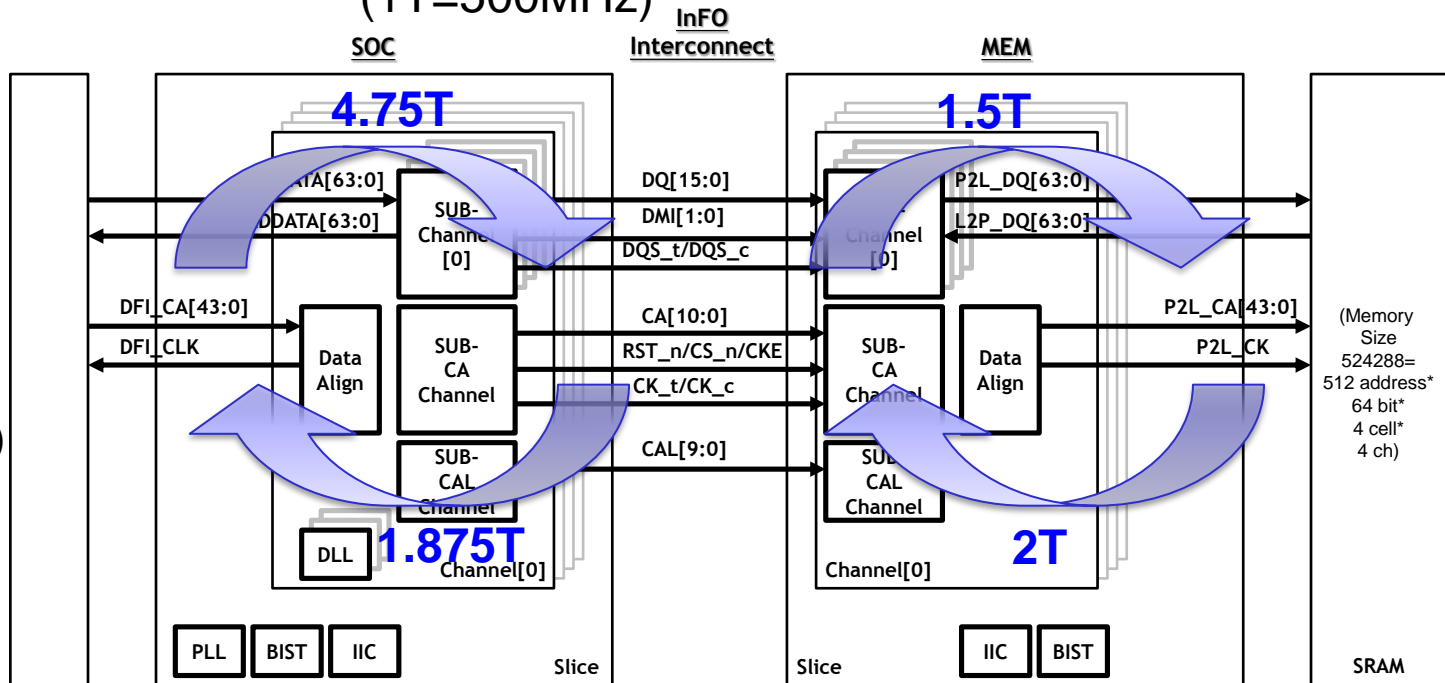
- ❑ 3 DLLs (TX/RX/RPT)
- ❑ Prompt and automatic
 - ❑ Lock time < 1us
- ❑ 10-pin overhead shared by 64-DQ
 - ❑ But scalable



17 Note: "RPT"="Read-Preamble-Training"

Latency-cost Degrades System Efficiency

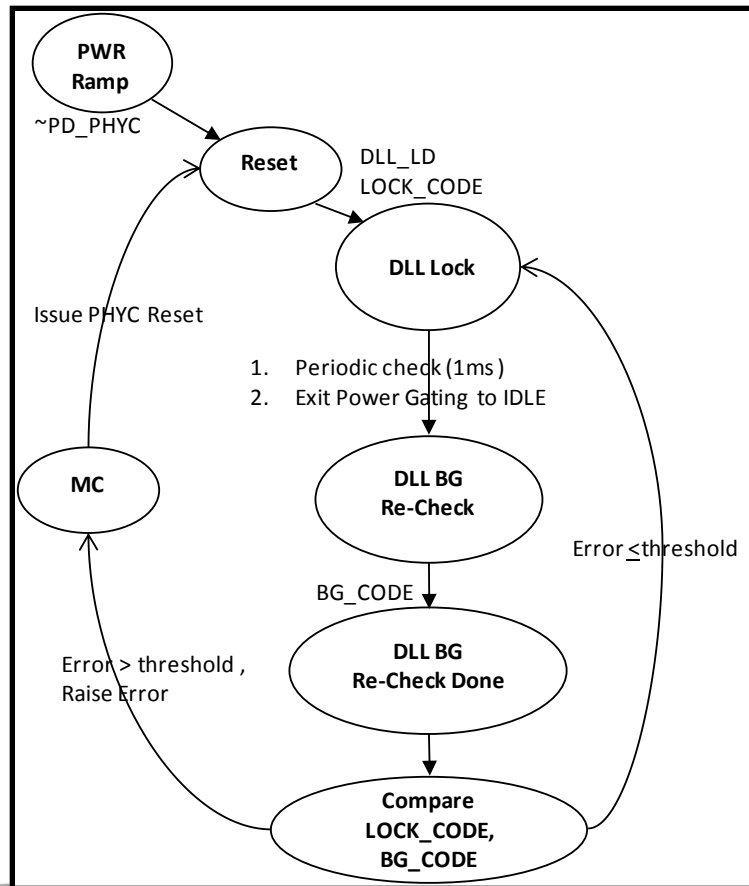
(1T=500MHz)



- ❑ De-/Serialization ratio: 4
- ❑ DBI enable (data-bus-inversion)

Temperature Drift Monitoring Scheme

- Temperature drift →
- Timing variation →
- Performance degradation
- Leveraging DLL architecture
- Periodic check (1ms) in background
- Alert issue if drift exceed the tolerance



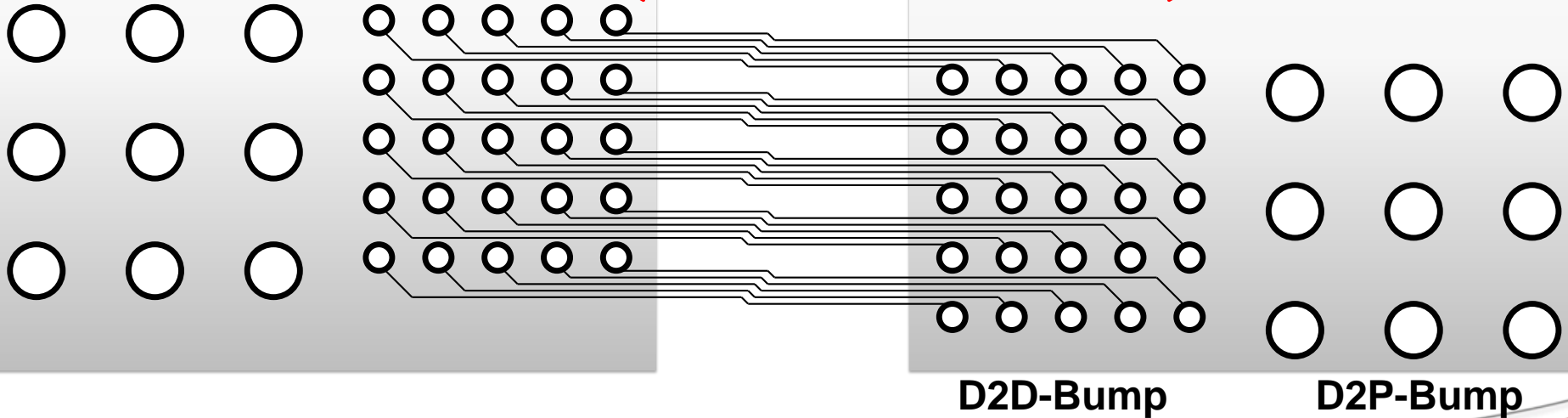
Pad Plan

- Compact pad plan with perfect-match trace length
- RDL routing density: $1 \times \text{RDL}/10\mu\text{m}$

D2P-Bump

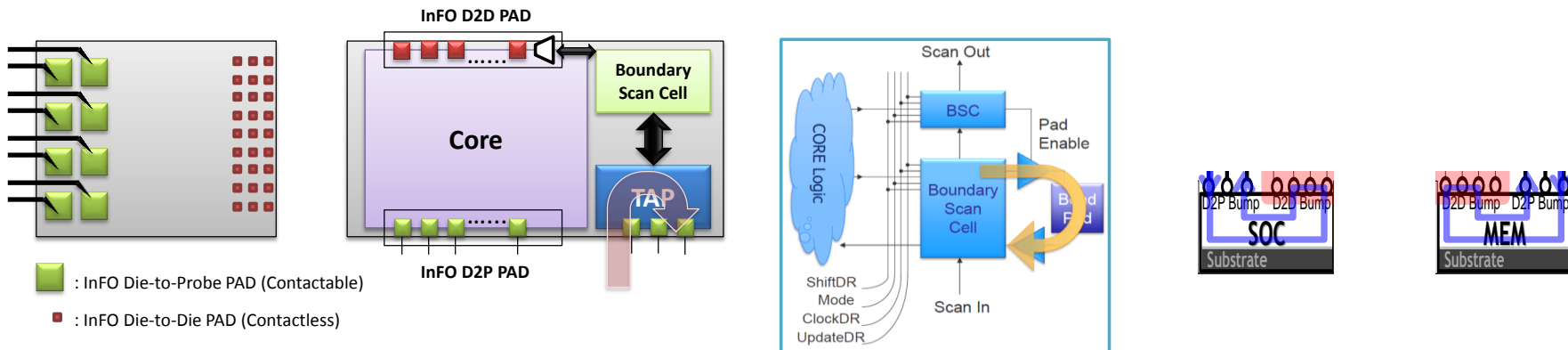
D2D-Bump

550um



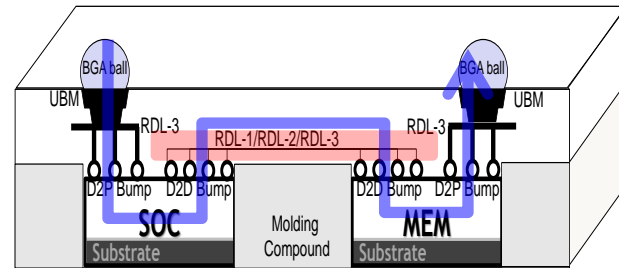
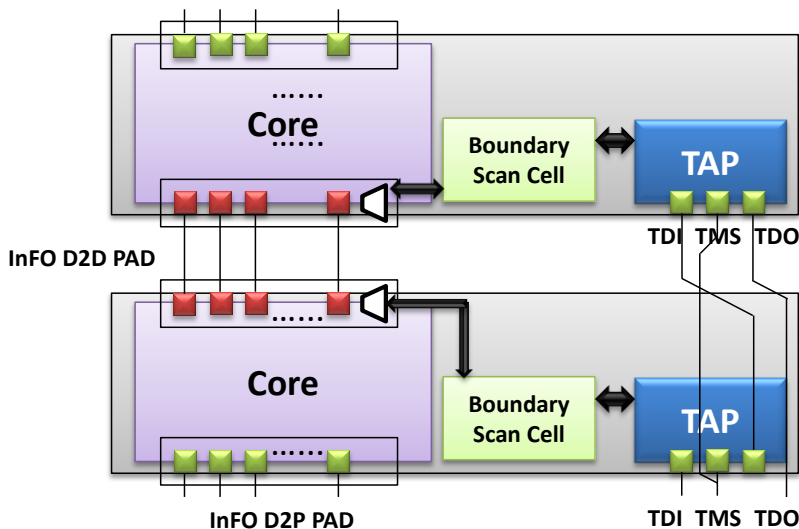
Boundary Scan Strategy

- **KGD (Known-Good-Die) – Contactless IO**
 - Compatible with **IEEE1149.1**
 - Each die builds in self TAP controller
 - Support contactless LIPINCON-IO open/short test individually



Boundary Scan Strategy

- **KGS (Known-Good-Stack) – Interconnect**
 - Cascade TAP structure and control in sequence
 - Support inter-connectivity test through TAP/Bscan cells
 - Capable of per pin diagnosis

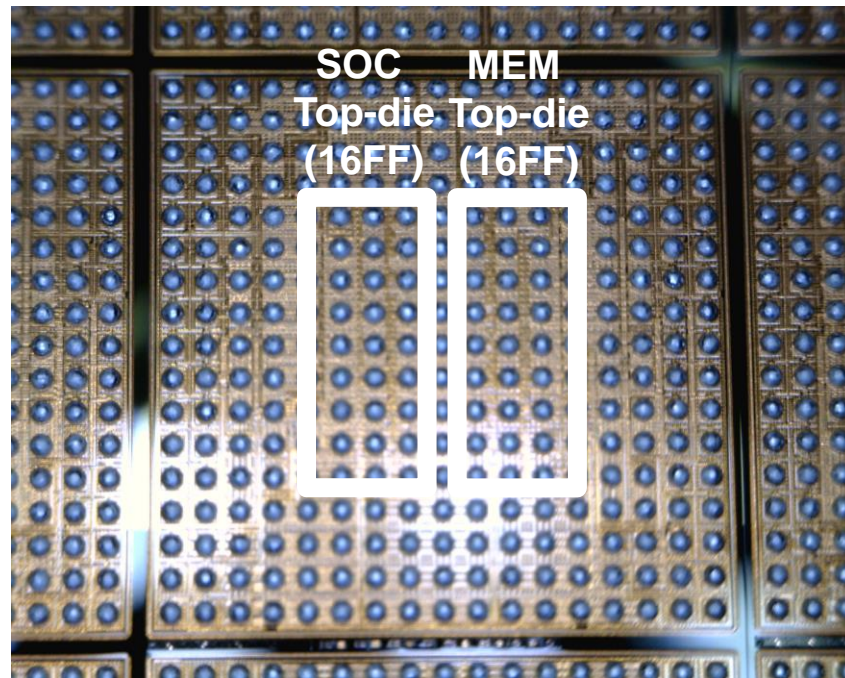
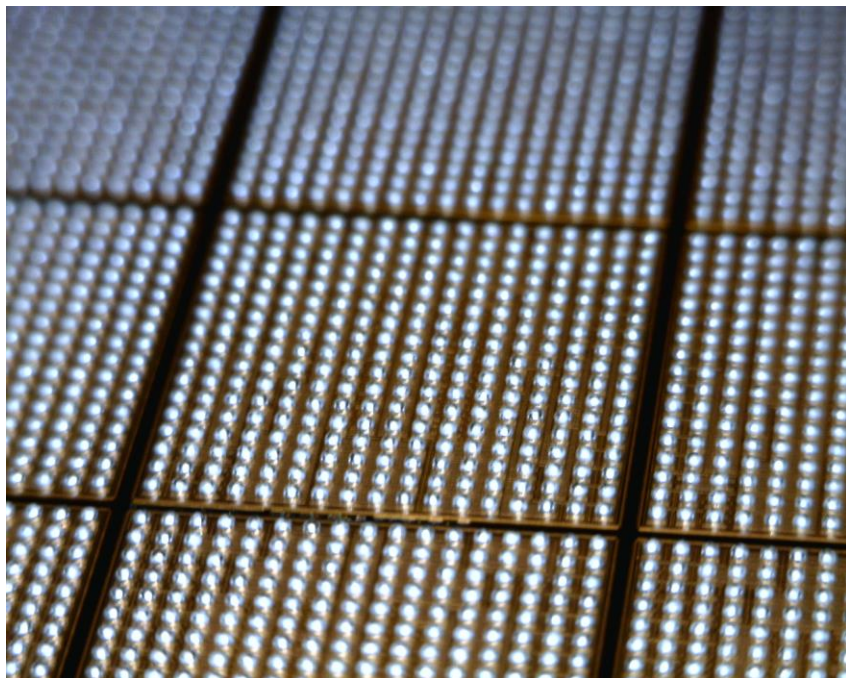


Packaged-Die Photo After InFO

- Fan-out to get the required direct-access BGA balls

[BGA ball side]

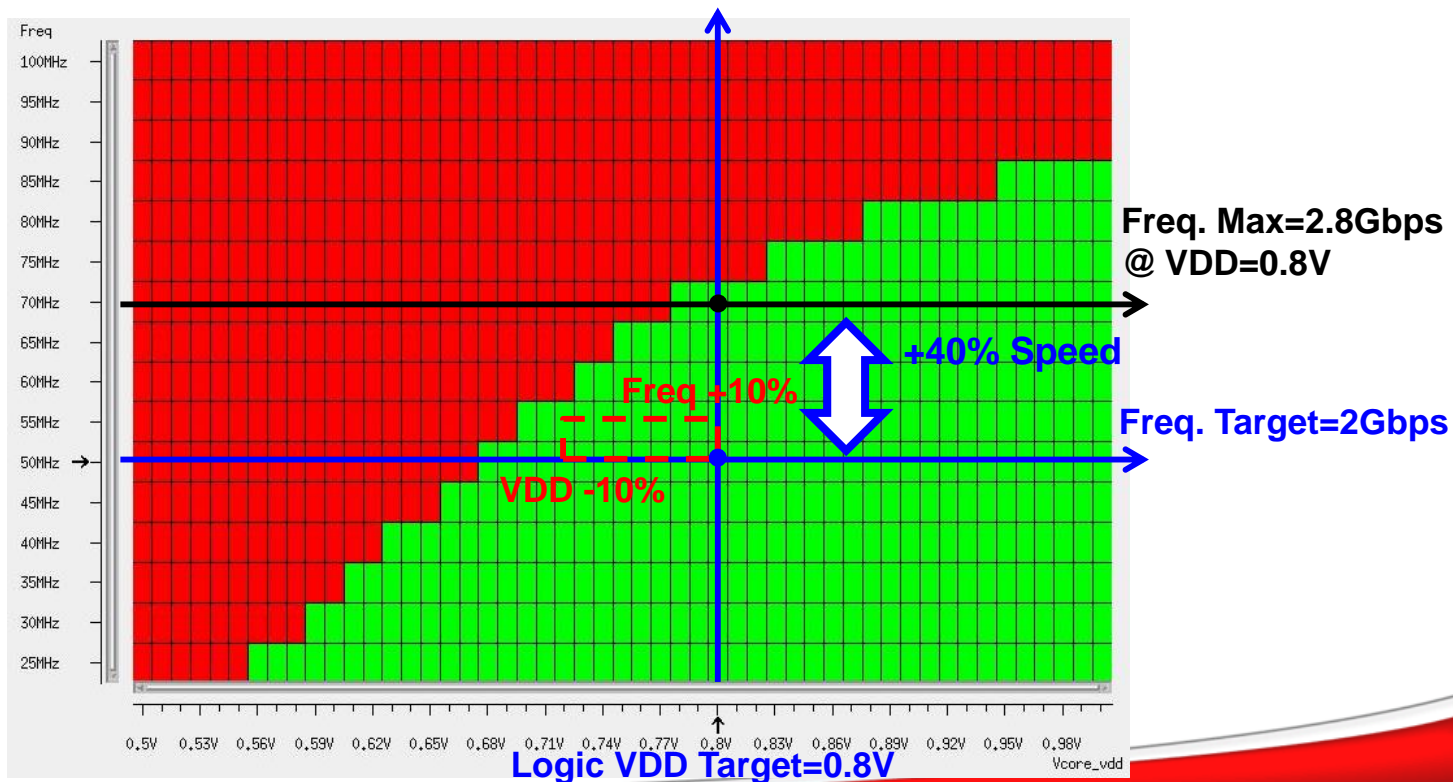
[7mm*7mm]



KGD Shmoo Plot

● SOC/MEM-PHY loopback BIST (Logic VDD vs. Frequency)

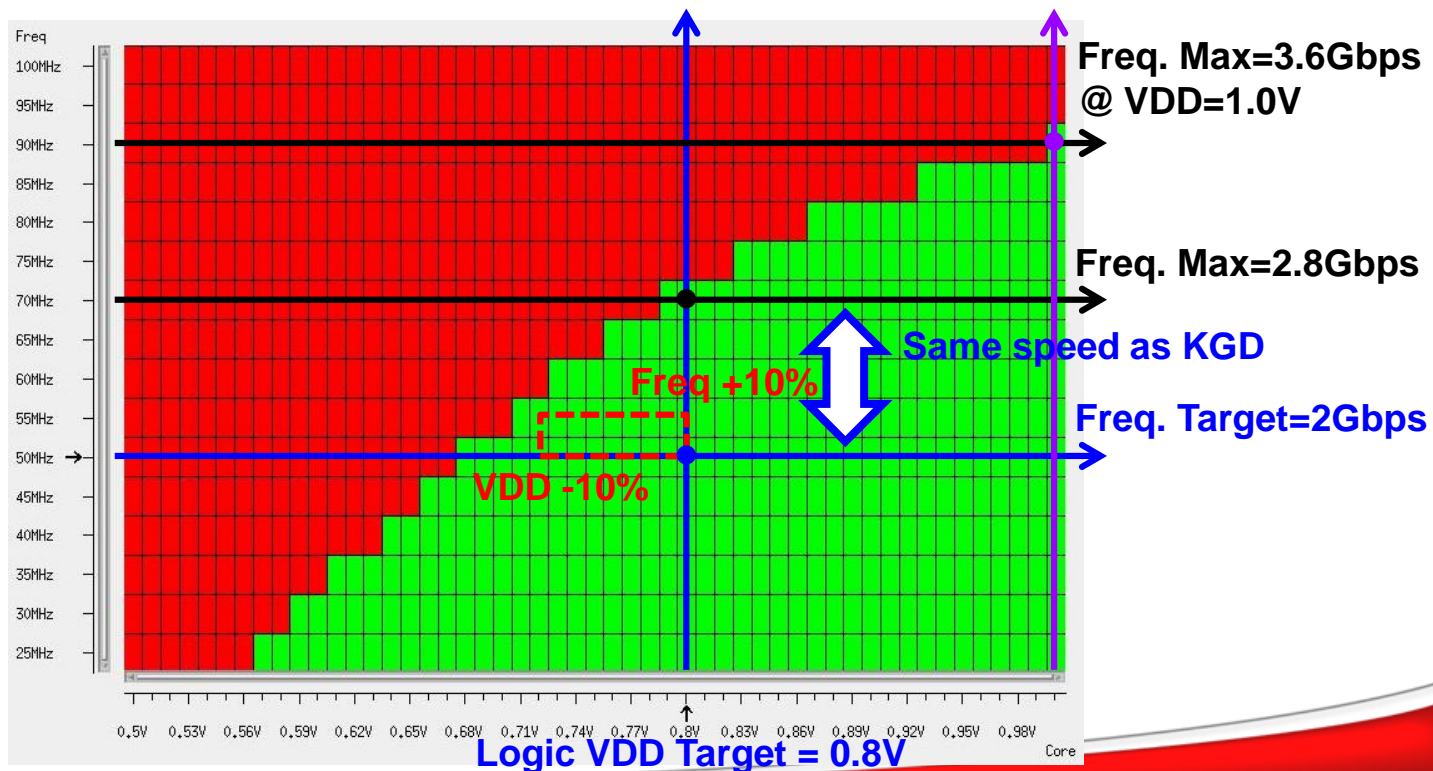
- VDDQ=0.3V
- 256-DQ toggle
- DBI enable
- PRBS



KGS Shmoo Plot

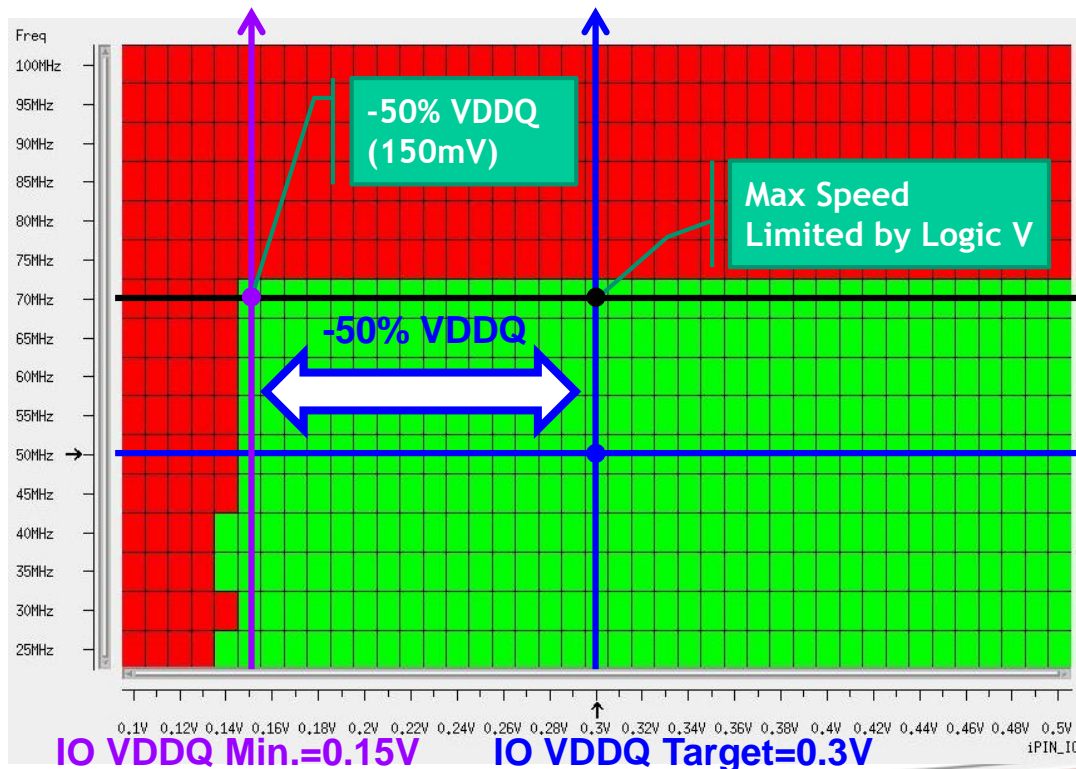
● SOC-to-MEM Write/Read BIST (Logic VDD vs. Frequency)

- VDDQ=0.3V
- 256-DQ toggle
- DBI enable
- PRBS



KGS Shmoo Plot

● SOC-to-MEM Write/Read BIST (IO-VDDQ vs. Frequency)

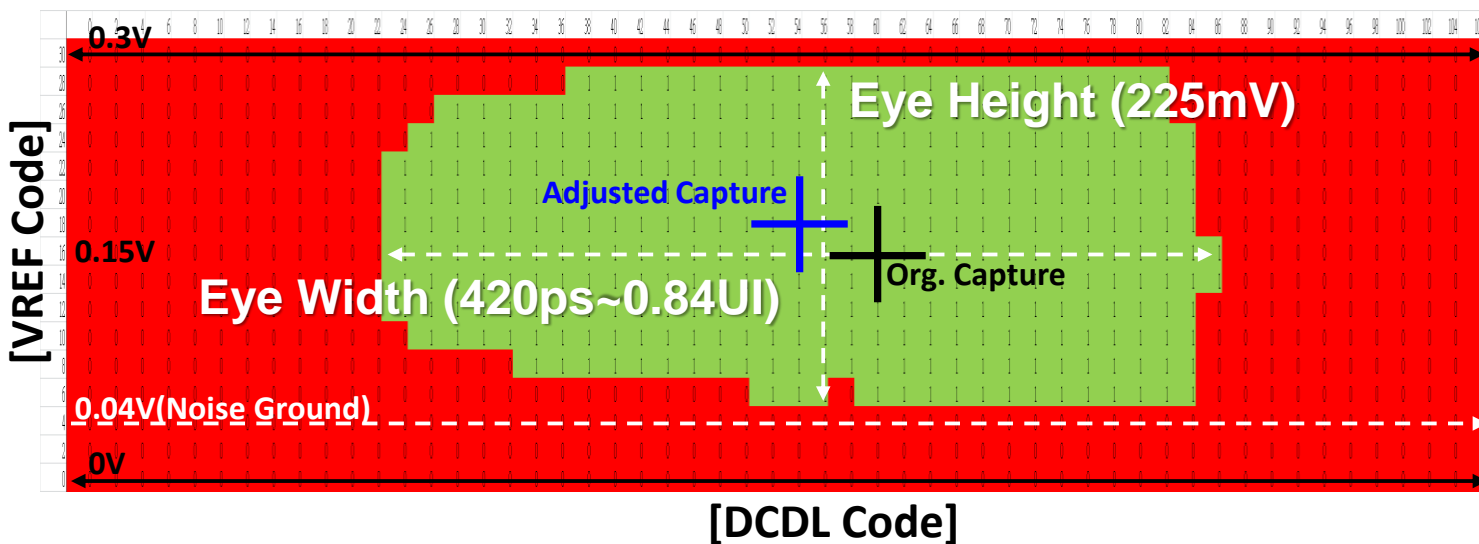


- VDD=0.8V
- 256-DQ toggle
- DBI enable
- PRBS

Capable of Eye-plotting

- 0.3V Swing is critical to SSO under wide bus application
 - How to ensure signal integrity on the un-probed interconnect IO?

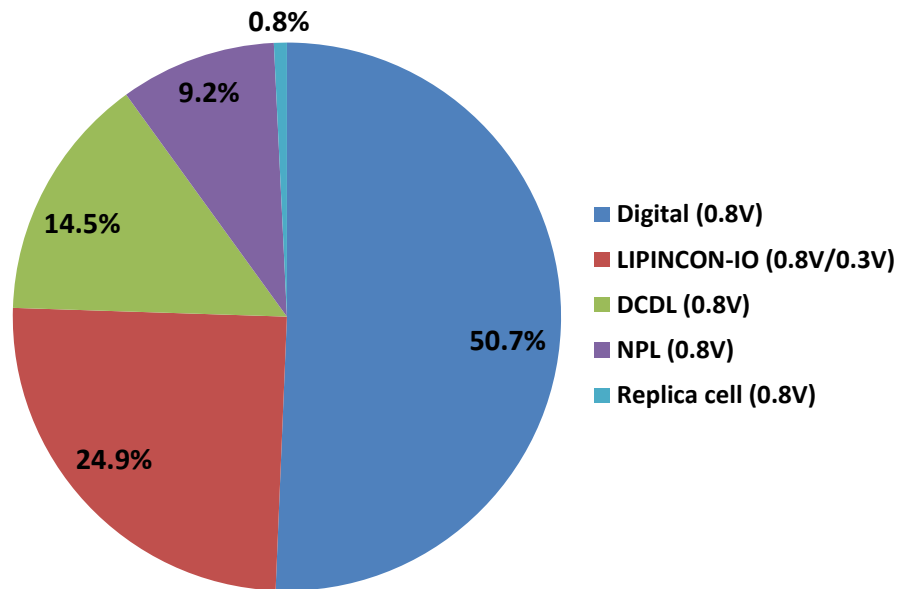
- ❑ X-axis: DCDL code (7ps)
- ❑ Y-axis: VREF code (9.5mV)
- ❑ VDD=0.8V
- ❑ 256-DQ toggle
- ❑ DBI enable
- ❑ PRBS



Power Breakdown

- Base on simulation results with 100% toggling conditions

Power Breakdown



- LIPINCON-IO power efficiency:
0.062pJ/bit (consider one single-end IO)
- LIPINCON-PHY power efficiency:
0.424pJ/bit

Blocks	Power Consumption	Power Efficiency	Percentage
Digital (0.8V)	110.00	0.215	50.7%
LIPINCON-IO (0.8V/0.3V)	53.97	0.105	24.9%
DCDL (0.8V)	31.52	0.062	14.5%
NPL (0.8V)	20.00	0.039	9.2%
Replica cell (0.8V)	1.65	0.003	0.8%
SUM	217.141	0.424	100.0%
	(mW)	(mW/Gb/s)	

Conclusion

- An in-package interconnect for in-package memory application in InFO package has been demonstrated
 - Technology: **TSMC 16FF + InFO**
- **89.6GByte/s** total bandwidth is achieved with **256-DQ** operating in **2.8Gbit/s** and **0.3V**-swing
 - Low power: **IO (0.062pJ/bit); PHY (0.424pJ/bit)**
 - Low latency: **Write (4.75T+1.5T=6.25T); Read (2+1.875=3.875T)**
- 0.3V signal integrity on the un-probed IO has been clarified
 - 420ps (**0.84UI**) Eye width; 225mV (**75%**) Eye height
- Prompt and automatic timing-calibration scheme

Acknowledgement

The authors would like to thank TSMC “**InFO-IP**” team for InFO back-end support, TSMC “**IPPM**” team for measurement support, and “**System-BD**” team for business help.

The authors would also like to specially thank “**Mentor Graphics**” for boundary scan design collaboration.

Reference

- [1] Intel, 2015-ISSCC, “The Xeon® Processor E5-2600 v3: A 22nm 18-Core Product Family”
- [2] Google, 2015-VLSI, “System Challenges and Hardware Requirements for Future Consumer Devices: From Wearable to ChromeBooks and Devices in-between”
- [3] TSMC, 2013-OIP “SoC Memory Interfaces. Today and tomorrow at TSMC”
- [4] 2012 Business Update “3DIC & TSV interconnects”
- [5] TSMC, 2012-IEDM, “High-Performance Integrated Fan-Out Wafer Level Packaging (InFO-WLP): Technology and System Integration”
- [6] TSMC, 2013-VLSI, “An extra low-power 1Tbit/s bandwidth PLL/DLL-less eDRAM PHY using 0.3V low-swing IO for 2.5D CoWoS application”
- [7] JESD8-28 “300mV interface”