Physical Design Challenges and Innovations to Meet Power, Speed, and Area Scaling Trend

LC LU
TSMC
TSMC Fellow/Senior Director, R&D
Chip and System Integration Trends for Better PPA & System Performance

- Better performance
- Better Power
- Better form factor

3D Transistors (FinFET, GAA-NWT)
3D Stacking (BSI, Monolithic MEMS)
3D Packaging (WLP, WoW)

Platform Solutions
- Mobile
- High Performance Computing
- Automotive
- IoT

SYSTEM INTEGRATION
- System performance/ Bandwidth
- Application-specific
- Process technology + Design Solutions

Multiple Chips on PCB

2D

SoC

3D

Open Innovation Platform®
Application-Optimized Design Platforms
Area, Performance and Power

- Speed & Memory Bandwidth
- Functional Safety
  Standard Compliance
- Ultra-Low Power
- Mobile
- Automotive
- IoT

High-Performance Computing
Semiconductor Trend

- **Area**: process primary dimension scaling slows down
  - Very challenging to continue Moore’s law economically
  - Process and design co-optimization to provide enough area scaling
- **Performance**: dimension scaling grows metal and via resistance exponentially
  - Fully automatic and smart via pillar design flow to reduce high resistance impact
- **Power**: ultra low voltage for ultra low power
  - Robust design and variation modeling at low voltages
- **Heterogeneous integration**
  - Low cost and high system performance 3D packaging
- **High physical design complexity**
  - Machine learning to be applied to physical design
Area Scaling

- Process primary dimension (metal, gate and fin pitch) scales less and per generation cadence takes longer
- Process-design co-optimization innovation to keep area scaling
  - Fin depopulation to increase cell density
  - Power plan and cell layout co-optimization to increase cell placement utilization
  - EUV to increase routing density
Fin Depopulation can Increase Density and Reduce Process Dimension Scaling Pressure

4-fin cell

3-fin cell

2-fin cell
Fin Depopulation can Increase Speed and Power Efficiency

- Higher fins have highest speed
- Fewer fins have highest speed @ same power and lowest power @ same speed
Fin Depopulation Impact on Logic Density and Cell Utilization

Logic Density

16nm 3~4 fins
10nm 3 fins
7nm 2 fins

Cell Utilization
80%
75%
70%
Logic Density Improvement (I)
Power Plan Optimization

- To improve power plan IR, PG via count is increasing over technology generations.

- However, shrinking PG pitch impacts routing resources. So different and new PG design approaches evolved.

  - Shrink PG Pitch to increase via count.

  - Allows better cell placement freedom.
Logic Density Improvement (II)
Power Plan and Cell Layout Co-optimization

Uniform M1

PG and Cell Optimized Co-Design

Dual M1

PG design has to be friendly to cell architecture (Uniform → Dual)

Cell re-Design

Big cell design has to be redesigned to be compatible with Dual PG architecture
Logic Density Improvement (III)
Power Plan and Cell Layout Push to Limit

Power strap:
Less cells placed under PG

Power stub:
More cells placed under PG

No stagger pin: 5 access points

Stagger pin: 6 access points

Unused space
Logic Density Improvement (IV)
EUV to Increase Routing Density

Inverse Litho.

Multiple patterning

EUV

Directed Self-Assembly

Hole type DSA, 16 nm half pitch

Line/Space DSA, 12 nm half pitch

NXE3300 EUV Single Patterning

6.6 nm
EUV Metal

Metal : Poly pitch = 2 : 3

- More Metal resources for logic density improvement
- 2 library sets are needed for two metal offset.

Metal : Poly = 1:1

Metal : Poly = 2:3
Less Coupling from Metal : Poly = 2 : 3

ND2D1
M1:PO pitch = 1:1

ND2D1
M1:PO pitch = 2:3
More Routing from Metal : Poly = 2 : 3

M1:Poly pitch = 1:1

M1:Poly pitch = 2:3
Enhanced Logic Density and Cell Utilization Trends

- Optimized power plan, pin access, EUV routing

Logic Density

- N16 11ML
- N10 12ML
- N7 13ML

Cell Utilization
- 80%
- 75%
- 70%
Performance Scaling

- Process dimension scaling grows metal and via resistance exponentially
- Fully automatic and smart EDA design flow is needed to effectively solve high resistance issue for 7nm and beyond
Cross Node Metal RC Scaling

- Increase ~3X
- Mx Resistor (Mx R) increases
- Mx Capacitor (Mx C) increases

Node Sizes:
- 40nm
- 28nm
- 16nm
- 7nm
- 5nm
Cross Node Wire R and Via Impact

BEOL R delay vs. Total delay

BEOL R Impact Increase

40nm  28nm  16nm  7nm  5nm

0%  10%  20%  30%  40%  50%
• Large drivers + upper thick metal + via pillar to reduce transistor, metal and via resistance
• VIA-Pillar enablement is complex and requires EDA enablement across all Implementation phases
Via Pillar Insertion

- Implies layer promotion for better wire resistance
- Effectively reduces source Via resistance

Connection through DPT layer: **DPT wire resistance dominates**

<table>
<thead>
<tr>
<th>Metal R</th>
<th>Source Via R</th>
</tr>
</thead>
<tbody>
<tr>
<td>55.43X</td>
<td>1X</td>
</tr>
</tbody>
</table>

Layer promotion to non-DPT layer: **Source Via resistance increases**

<table>
<thead>
<tr>
<th>Metal R</th>
<th>Source Via R</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.23X</td>
<td>3.29X</td>
</tr>
</tbody>
</table>

Via pillar: **Best combination of metal and source Via resistance**

<table>
<thead>
<tr>
<th>Metal R</th>
<th>Source Via R</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.23X</td>
<td>1X</td>
</tr>
</tbody>
</table>
Cross Node Wire R and Via Impact

BEOL R delay vs. Total delay

BEOL R Impact Increase

with Via Pillar

40nm  28nm  16nm  7nm  5nm
Via Pillar Type and Usage

- EM via pillar (EM VP), used to guarantee cell-level EM, has to be 100% inserted
  - One cell master will only have 1 EM VP
- Performance via pillar (Performance VP), which is larger than EM VP, can reduce more resistance
  - One cell master can have multiple performance VP

**VP Association I**

**Cell Master** → Performance VP1 → Performance VP2 → Performance VP.. → EM VP → Default VP

**VP Association II**

**Cell Master** → EM VP → Performance VP1 → Performance VP2 → Performance VP..
Via Pillar Automatic Design Flow

- **Design Kits**
  - VIA pillar (VP) definition, association

- **Floorplan**

- **Placement**
  - VP aware placement
  - Automatic NDR

- **CTS**
  - VP insertion and verification
  - CTS with VP

- **Routing/Opt.**
  - VP optimization
  - DEF out option
Voltage and Power Scaling

- Low voltage design reduces power effectively
- Low voltage design challenges
  - Functionality robustness
  - Accurate variation modeling
Ultra-Low-Voltage Standard Cell Solution

- Skew Cells/ D0 or Fine-Grained Cells
- Internal Half-Sizing Cells
- Small High-Stack(3 or 4) Cells
- Transmission Gates
- Multi-Bit Latch/Flop

0.9V-0.6V

- Proper Stage Ratio to Maintain Internal Slew
- Proper P/N Ratio to Improve Noise Margin
- New Circuit Structure to Enhance Robustness at Low-Vdd

0.5V and below

- Improve cell performance by reducing delay degradation and variation induced by low-Vdd
Flop Low Voltage Design Robustness

- High sigma design checks are required to ensure robust operation at low operating voltage
- For write operation, forward path (red) must be stronger than feedback path (blue)
Delay Variation Increases at Lower VDD

Delay variation increases exponentially as VDD decreases.

Non-Gaussian at low voltage

Gaussian at regular voltage
Abnormal Hold-Time Caused by Low-Vdd Non-Gaussian Behavior

Flop Hold Constraint Behavior

Regular VDD

Low VDD
Solutions for Low-Vdd Variations (I)

- Methodology improvement
  - Timing values are models by both “mean” and “early/late distribution”
  - Implement the new models in timing characterization and STA tools
Solutions for Low-Vdd Variations (II)

- By using new timing model and advanced statistical OCV method, we achieve more accurate STA results compared to Monte-Carlo simulations.
Heterogeneous Integration

- Low cost and high system performance
- InFO and CoWoS integration
- Integrated EDA design flow for package and chips
Future Device Possibilities Through Heterogeneous Technology Integration

3D Low Power CMOS

Integrated Specialty Technology

3D-IC Technology

- GAA- NWT
- TFET
- III-V FFET
- Si CMOS
- Ge FinFET on Si
- FinFET

- RF CMOS
- eFlash
- BSI CIS
- Mixed Signal
- Analog

- InFO
- CoWoS™ Si Interposer
- Cu-TSV Vertical Stacking
- BCD - Power IC
- CMOS MEMS

ISPD 2017
3D Packaging for Better System Performance and Form Factor

Conventional SIP or MCM (Wirebond, Flipchip, SMT)

3D Wafer-based System Integration (CoWoS, InFo-PoP, Vertical Stacking, Wafer Bonding)

Sources: A-SSCC 2014, IEDM 2014, VLSI 2015
ISPD 2017
Different Systems Require Different Packaging Solutions

- **WLCSP**: Wafer Level Chip Scale Package
- **InFO**: Integrated Fan-Out
- **CoWoS**: Chip on Wafer on Substrate

* I/O Pin Count

* Die/PKG size (mm²)
Wafer-Level Packaging Technologies: Integrated Fan-Out (InFO)

Multi-chips integration
Smallest form-factor
Cost competitive

- InFO
- InFO PoP
- InFO-M

WLCSP*

Die/PKG size (mm²)

I/O Pin Count

0
1000
2000
3000
4000

Logic
DRAM
Memory

Integrated Fan-Out (InFO)

InFO PoP

InFO-M (Multi-chip)

Through-InFO-Via (TIV)

WLCSP*
Wafer-Level Packaging Technologies: Chip-on-Wafer-on-Substrate (CoWoS)

- Homogeneous partition
- High bandwidth memory integration for high-performance computing
- High-performance heterogeneous partitioning

CoWoS®

I/O Pin Count

Die/PKG size (mm²)
Advanced Packaging Co-Design – SoC + InFO + IPD (Integrated Passive Device)

FC-PoP

<table>
<thead>
<tr>
<th>L6</th>
<th>L5</th>
<th>L4</th>
<th>L3</th>
<th>L2</th>
<th>L1</th>
<th>TMV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Silicon</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Molding Compounds</td>
</tr>
</tbody>
</table>

$R_{ja} (17.2 \, ^{\circ}C/W)$

590μm

InFO-PoP

<table>
<thead>
<tr>
<th>F-S pad</th>
<th>RDL3</th>
<th>RDL2</th>
<th>RDL1</th>
<th>RDL0</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM4</td>
<td>PM3</td>
<td>PM2</td>
<td>PM1</td>
<td>PM0</td>
</tr>
</tbody>
</table>

$R_{ja} (15.4 \, ^{\circ}C/W)$

250μm

1. Thermal dissipation improved by 12%
2. Thickness reduced by 57%
3. Eye width improved by >12%
4. Voltage droop reduced by 5–10%

Eye Opening: 0.785 UI → 0.663 UI

Eye Opening: 0.663 UI

Eye Opening: 0.785 UI

InFO-PoP W/I IPD

FC-PoP W/O IPD

ISA 2017

Open Innovation Platform
InFO Design Solutions (InFO Only)

PKG Layout In-design DRC

InFO Layout Creation

CDL Netlist Auto-Export
- PKG Layout
- InFO GDS
- InFO CDL Netlist
- InFO LVS

InFO DRC/LVS

RLCK Extraction
< 4GHz RC/RLCK Extraction for STA, IR, SEM, & PI Analysis

InFO RC Extraction

S-parameter Extraction
4GHz – 10GHz S-parameters Extraction for SI/PI, & EMI analysis

ISPD 2017
InFO Design Solutions
(Dies + InFO)

Inter-die DRC/LVS

Inter-die DRC

Inter-die LVS

Top die

EM/IR Analysis

Static IR Analysis

Dynamic IR Analysis

SI/PI Simulation

SI Analysis

PI Analysis

SI/PI Simulation

Thermal Analysis

Thermal-aware EM/IR Analysis

Normal Analysis (No Thermal)

Thermal-aware Analysis

Open Innovation Platform®
Applying Machine Learning to Complex Physical Design Problems

- Place and route Machine Learning experiment
- Feature extraction and convolutional neural network data mapping
- Clock gating and routing congestion speed improvement
Machine Learning-Based Design Solution

- Eliminate human & fixed-model subjective bias with statistical important features

### Traditional EDA Approach

**Prediction**
- Placement result
- Routing pattern model

**Routing simulation programming**

**Routing overheads**

Easy to be biased and no guaranteed quality

Only heuristics are possible

### Machine Learning Approach

**Learning**
- Placement result
- Detailed routing result

**Model training**

**Routing overheads**

Leverage widely used deep learning packages

### Traditional (EDA) flow

**Pre-route design**

**Congestion map after routing**

Only know routing results after running routing

### Flow enables by Machine Learning

**Congestion prediction**

**Congestion map after routing**

Able to predict routing behavior and improve congestion with new recipes to achieve 40Mhz speed improvement
ML Design Enablement Platform

Adopt TSMC ML Platform to predict the design improvement opportunity

TSMC ML Design Enablement Platform

- Feature extractions
- Data pre-processing
- Prediction

Script to perform design optimization

TSMC’s proven ML training models
ML Design Enablement Platform

Adopt TSMC ML Platform to train design and build new models

APR database for new learning

TSMC ML Design Enablement Platform

- Feature extractions
- Data pre-processing
- Training
- Model Mgmt.

Prediction

- Script to perform design optimization

- TSMC’s models
- Customer’s models

Tapeout

Optimized design

RTL → Synthesis → Place & Route → STA → Signoff → Tapeout

ISPD 2017
Performance Gains Prediction from ML

- Capability to predict and set accurate ARM A72 clock gating latency to avoid over-design and achieve better speed from 50 to 150Mhz
- Post-route speed is improved by 40MHz with detour prediction and early fix

Achieved 95% matching on A72
Conclusion

- Five semiconductor industry trends, issues and solutions are discussed: area, performance and power scaling, heterogeneous 3D integration and Machine Learning
- Physical design and EDA play even more critical roles to extend Moore’s law and to enable highly integrated complex 3D SOC chips