Nanometer IC Design for Manufacturability and Reliability in Extreme Scaling and Beyond

David Z. Pan
Dept. of Electrical and Computer Engineering
The University of Texas at Austin
http://www.cerc.utexas.edu/utda
Nanometer Issues
The industry forced to extend 193nm lithography

- Immersion, extreme RET
- Double/Triple Patterning Lithography for 22nm/14nm, ...
Resolution Limit of Optical Lithography

With single exposure, the practical limit is ~80 nm

$$\text{pitch} > \frac{0.5\lambda}{n \sin \theta}$$

$$= \frac{0.5 \times 193}{1.35}$$

$$= 72 \text{ nm}$$
What is Double Patterning?

Overlay Error

Stitch

\( \text{min}_{dp} \)
Back in 2003

Jeffrey: "Daddy, don’t worry. I can show you some bright ideas!"

“Double patterning!!!”
Emerging Lithography

193i w/ DPL

Quadruple patterning

EUV

DSA

E-beam

3D-IC
The “Moore”, the Merrier!

♦ More Moore
  • Nano-Patterning for Extreme Scaling
  • Lithography Aware Physical Design

♦ A different kind of “Moore”
  • 3D Integration
  • New devices (e.g., nanophotonics) …

➡ Need synergistic design and technology co-optimization for cross-layer resilience
DPL Layout Decomposition

Minimum Stitch Insertion

1) Minimize stitch #

2) A bit more overlap margin for stitch, but area increases

[Lucas SPIE’08]

Overlay Compensation

Without Overlay Compensation

With Overlay Compensation

[Yang+, ASPDAC10]
A Graph-Partitioning Based, Multi-Objective Decomposer

Decomposition Graph Construction

[Yang+, ASPDAC10]

Constraint: 
(A, Ā) and (E, Ė) are repulsive pairs.

Theorem: Stitch minimization problem is equivalent to the min-cut partitioning of the decomposition graph

Extensions of the framework: to incorporate other constraints and costs into graph partitioning, e.g., balanced density, overlay compensation, and so on
Overlay Compensation & Density Balancing

[C432: 27% and 73% (7 stitches)]

[C432: 50% and 50% (17 stitches)]

[Without TDD: 9% variation]
[One stitch: 5.274% variation]
[Three stitches: 1.098% variation]
[Nine stitches: 0.018% variation]

[Yang+, ASPDAC10]
Triple Patterning

- Extension of double patterning concept
- Original layout is divided into three masks
  - Resolve native conflicts of DPL
  - Achieve further feature-size scaling (14nm, 11nm)

- Very active research area!
- DAC’12 Best Paper; ICCAD’13 Best Paper
- ICCAD’13 Special Session
  - Session 3B: Triple Patterning, Triple the Trouble?
Electronic Beam Lithography

- Maskless technology, which shoots desired patterns directly into a silicon wafer
- Low throughput is its major hurdle
  - E.g., Variable Shaped Beam (VSB)

Total number of 11 shots are needed
Character Projection (CP) Technology

- Print some complex shapes in one electronic beam shot, rather than writing multiple rectangles.

Diagram showing the process of shaping characters using an electron gun, stencil, and shaping aperture, with the final result showing three shots only.
Overlapped Characters Planning
[Yuan and Pan, ISPD’11 Best Paper Award]

♦ The number of characters is limited due to the area constraints of the stencil

![Diagram showing character layouts and spacing](https://example.com/diagram.png)

♦ By over-lapping adjacent characters, more characters are allowed to be put on the stencil
The “Moore”, the Merrier!

♦ More Moore
  • Nano-Patterning for Extreme Scaling
  • Lithography Aware Physical Design

♦ A different kind of “Moore”
  • 3D Integration
  • New devices (e.g., nanophotonics) …

➔ Need synergistic design and technology co-optimization for cross-layer resilience
LAPD

- Double/multiple patterning layout compliance/decomposition
- Post-layout, maybe too late!
- Lithography Aware Physical Design (LAPD)
- Litho Hotspot Detection
- Litho Friendly Design
  - Hotspot Avoiding/Correction
  - Correct by Construction/Prescription
Lithography Hotspot Detection

- Lithographic hotspots
  - What you see (at design) is NOT what you get (at fab)
  - Hotspots mean poor printability
  - Highly dependent on manufacturing conditions
  - Exist after resolution enhancement techniques

- Litho-simulations are extremely CPU intensive
  - Full-blown OPC could take a week
  - Impossible to be used in inner design loop
Various Approaches

**Pattern/Graph Matching**
- Pros and cons
  - Accurate and fast for known patterns
  - But too many possible patterns to enumerate
  - Sensitive to changing manufacturing conditions
  - High false-alarms

**Data Mining/Machine Learning**
- SVM [J. Wuu+ SPIE09]
  [Drmanac+ DAC09]
- Neural Network Model
  [Norimasa+ SPIE07][Ding + ICICDT09]
- Regression Model
  [Torres+ SPIE09]
- Pros and cons
  - Good to detect unknown or unseen hotspots
  - Accuracy may not be good for “seen” patterns (cf. PM)
  - Hard to trade-off accuracy and false alarms

[Xu+ ICCAD07]
[Yao+ ICCAD08, [Khang SPIE06], etc.]
A New Meta-Classification Paradigm

Pattern Matching Methods
Good for detecting previously known types of hotspots

Machine Learning Methods
Good for detecting new/previously unknown hotspots

A New Unified Formulation (EPIC)
Good for detecting all types of hotspots with advantageous accuracy/false-alarm (Meta-Classifier)

Meta-Classification combines the strength of different types of hotspot detection techniques

[Ding et al, ASPDAC 2012]
Components of Meta-Classifier Core

- Base classifier results are first collected
- Weighting functions to make the overall meta decision (e.g., quadratic programming)
- Threshold with accuracy and false-alarm trade-off
Using the machine learning models, we built a new detailed router AENEID to avoid hotspot patterns.
Abstraction to Synthesis & Above?

- Can we further extend the abstraction up to logic synthesis?
  - Not just lithography hotspot, but other hotspots such as reliability metrics including BTI, oxide breakdown
- Machine learning to raise the abstraction?
- NSF/SRC FRS program (started April 1, 2003)
- NSF/SRC/DFG Cross-Layer Resilience Workshop in Austin, July 11 and 12
The “Moore”, the Merrier!

♦ More Moore
  • Nano-Patterning for Extreme Scaling
  • Lithography Aware Physical Design

♦ A different kind of “Moore”
  • 3D Integration
  • New devices (e.g., nanophotonics) …

⇒ Need synergistic design and technology co-optimization for cross-layer resilience
Thermal/Mechanical Stress

CTE : Coefficient of thermal expansion

<table>
<thead>
<tr>
<th>Material</th>
<th>CTE in $10^{-6}$/K at 20°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>3</td>
</tr>
<tr>
<td>W</td>
<td>4.5</td>
</tr>
<tr>
<td>Cu</td>
<td>17</td>
</tr>
</tbody>
</table>

TSV: 250 °C ~400 °C process (Higher than operating temperature)
Since Cu has larger CTE than Si $\Rightarrow$ tensile stress in Si near TSV.
Stress => Variability/Reliability

- Systematic Variations
  - Mobility
  - Timing

- Reliability (interfacial crack, EM, etc.)

Interfacial Crack

Electromigration Effect – Open

Electromigration Effect – Short

Image of interface crack and other illustrations.
Full-chip stress analysis considering multiple TSVs

TSV1, TSV2, and TSV3 affect P

TSV4 doesn’t affect P

stress influence zone (25um)

P: point under consideration
(Lateral &) Vertical Superposition

- Stress components are added up “vertically”

LVLS works well
# Reliability/Variability Impact of Stress

## 1. Von Mises Reliability

- Von Mises Yield is function of stress tensor

![Von Mises stress with TSV array](image1)

![Von Mises stress with three TSVs](image2)

[J. Mitra et al., ECTC’11]

## 2. Crack: Energy release rate (ERR)

- TSV stress affects ERR of TSV structure → aggravate crack

![Side view](image3)

![Top view](image4)

[M. Jung et al., ICCAD’11]

## 3. Mobility/ $V_{th}$ variation of MOS

- TSV stress changes mobility of hole/electron → timing, $V_{th}$ variation

![Hole mobility variation](image5)

![Electron mobility variation](image6)

[J. Yang et al., DAC’10]
From Stress to Reliability

♦ Von Mises Reliability Metric

\[
\sigma_v = \sqrt{\frac{(\sigma_{xx} - \sigma_{yy})^2 + (\sigma_{yy} - \sigma_{zz})^2 + (\sigma_{zz} - \sigma_{xx})^2 + 6(\sigma_{xy}^2 + \sigma_{yz}^2 + \sigma_{zx}^2)}{2}}
\]

♦ Physical meaning

If \( \sigma_v > \) yielding strength, deformation will be permanent and non-reversible.

Yielding strength
- Cu: 225 ~ 600 MPa
- Si: 7,000 MPa
Wide I/O 3D DRAM

(a) Pkg-bumps are placed underneath TSV arrays
(b) Pkg-bumps are placed 200μm apart from TSV arrays

<table>
<thead>
<tr>
<th>case</th>
<th>von Mises stress distribution (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>780-810</td>
</tr>
<tr>
<td>(a)</td>
<td>30</td>
</tr>
<tr>
<td>(b)</td>
<td>182</td>
</tr>
</tbody>
</table>

Case (b) shows that chip/package co-design can greatly reduce mechanical reliability problem in TSV-based 3D ICs

[Jung et al, DAC’12]
Cu shrinks faster than Si under negative thermal load ($\Delta T = -250^\circ C$)
- Model through Energy Release Rate (ERR)
- Full chip model with design-of-experiments of different layout styles and multiple TSV structures

[Jung et al, ICCAD’11]
Full-Chip Crack Analysis and Study

♦ Regular vs. irregular TSV arrays

(a) IrregA
(b) RegA

(c) ERR map of IrregA
(d) ERR map of RegA

100um

ERR (J/m²)

2.15
2.10
2.05
2.00
1.95
1.90
1.95 ~ 2.00
1.90 ~ 1.95
1.85
1.80
Stress Effect on Mobility & Current

CMOS (Stress: 200MPa, R=r) [Yang+, DAC’ 10]

NMOS: 0.5 $\Delta \mu$ ($\Delta I_{ds}$:+1.5%)
PMOS: 0.6$\Delta \mu$ ($\Delta I_{ds}$:+1.8%)

NMOS: 0.75$\Delta \mu$ ($\Delta I_{ds}$:+2.25%)
PMOS: -0.1$\Delta \mu$($\Delta I_{ds}$:-0.3%)

Cell characterizations based on distance and orientation are needed
Stress Aware Design Flow [Yang+, DAC’10]

1. Pre-placed TSV location → Stress estimation induced by TSVs
2. Stress estimation induced by TSVs → Mobility change ($\Delta \mu/\mu$) calculation
3. Mobility change ($\Delta \mu/\mu$) calculation → Cell characterization with mobility (Cell name change in Verilog)
4. Cell characterization with mobility (Cell name change in Verilog) → Stress aware Verilog netlist
5. Stress aware Verilog netlist → Verilog, SPEF merging for 3D STA
6. Verilog, SPEF merging for 3D STA → 3D Timing Analysis with PrimeTime
7. 3D Timing Analysis with PrimeTime → Critical gate selection
8. Critical gate selection → TSV stress aware layout optimization
9. TSV stress aware layout optimization → Optimized layout with TSV stress
10. Optimized layout with TSV stress → Liberty file having cell timing with different mobility
11. Liberty file having cell timing with different mobility → Verilog netlist
12. Verilog netlist → Stress estimation induced by TSVs
13. Stress estimation induced by TSVs → Mobility change ($\Delta \mu/\mu$) calculation

Stress-Aware ECO

Original cell placement

After cell perturbation

Rising critical optimization with hole contour

Falling critical optimization with electron contour
The “Moore”, the Merrier!

♦ More Moore
  • Nano-Patterning for Extreme Scaling
  • Lithography Aware Physical Design

♦ A different kind of “Moore”
  • 3D Integration
  • New devices (e.g., nanophotonics) …

➡ Need synergistic design and technology co-optimization for cross-layer resilience
Nanophotonics On-chip Integration

Holistic Optical Interconnect Planning and Synthesis

- Co-design and optimization with electrical interconnect
- Optical interconnect library (OIL) [Ding+, DAC’09, SLIP’09, and available http://www.cerc.utexas.edu/~ding/oil.htm]
- WDM, partitioning, routing, …

Nanophotonics is a very active field

Many new research problems for CAD community!
Case Study 1: O-Router [Ding et al, DAC’09]

- wafer-to-wafer TSV
- optical and electrical data conversion

(a) Metal Layers
(b) Metal Layers
(c) Metal Layers

Objectives: performance (throughput, latency, power), cost ($$, economics)

Constraints (SNR, signal integrity, reliability, system-level reqs.)
Case Study 2: GLOW

- Global router for low-power thermal-reliable optical interconnect synthesis using Wavelength Division Multiplexing (WDM)

Minimize Power (Laser driving power)

Synthesis Engine GLOW for On-Chip Optical/Electrical Interconnect

Circuit implementation

Max{inter. Delay} <= A
Max{SNR} <= B
Min{Ther. Reliability} >= C
Min{Pathend power} >= D

[Ding et al, ASPDAC’12]
Conclusion

- Optical lithography still pushing ahead for 14nm, 11nm, 7nm ➔ extreme scaling
  - Multiple patterning, EUV, DSA, and hybrid lithography
- Design enablement with lithography capability co-optimization from mask to physical synthesis (and higher level, e.g., NSF/SRC FRS)
  - Cross-layer resilience
- Horizontal scaling ➔ Vertical scaling: 3D-IC
  - Reliability/Variability issues
- New material/devices ➔ new CAD paradigms and tools
Acknowledgment

Support/collaboration from NSF, SRC, NSFC, Sematch, IBM, Intel, Oracle, Fujitsu, Qualcomm, Synopsys, Mentor Graphics, Toshiba, ...

The materials in this talk include results from many former/current PhD students at UTDA
  ‣ They are the ones who did the real work!

Collaborators
  ‣ DFM: IBM, Intel, Globalfoundries, Mentor, Synopsys, etc.
  ‣ 3D-IC: Prof. Sung Kyu Lim’s group at Georgia Tech
  ‣ Optical: Prof. Ray Chen’s group at UT Austin