Die-to-Package Coupling Extraction for Fan-Out Wafer-Level-Packaging

Yarui Peng¹, Dusan Petranovic², and Sung Kyu Lim³
¹University of Arkansas, Fayetteville, AR
²Mentor Graphics, Fremont, CA
³Georgia Institute of Technology, Atlanta, GA
Introduction

- Fan-out wafer-level-packaging (FOWLP)
  - Low-cost multi-die solution with excellent thermal and RF properties
  - Allows multiple chip to be tightly integrated with package
  - New parasitics are introduced in the interface layers
TV1: Mold-first for Fine RDL

- Mold frame
- Chip placement on molding tape on mold frame
- Wafer level molding
- Release from mold frame
- RDL and bumping processing
- Singulation

TV2: RDL-first for Fine RDL

- Formation of RDL and UBM on Carrier
- Die-to-Wafer Bonding
- Wafer Molding
- Support Carrier Removal & bumping

Die-to-Package Coupling Extraction for Fan-Out Wafer-Level-Packaging
Package chip dimension will keep scaling into 1um range using foundry BEOL process

- Package layers becomes similar to chip top metal layers

- Multi dies, SoC partition, 3D compatible
- < 2um L/S by foundry interposer BEOL
- High performance (CPU/GPU), mobile AP, BB

- Single/Multi die, SoC partition, 3D
- 2~10um L/S interconnect by bumping RDL
- Mobile AP / BB, mid-range CPU/GPU

- Single die
- RDL > 10um L/S
- RF, WLAN, Power etc.
Apple A10, A11 processor in iPhones all used FOWLP

- High performance, low cost, high IO density, very thin package
**Motivation**

- **Pkg-to-die (P2D) coupling noises threaten signal integrity:**
  - Significantly higher noise when both L and M elements are considered

- No parasitic extraction tool exists for P2D extraction
Our FOWLP Process Design Kit

- We designed a FOWLP PDK:
  - Chip: 45nm technology with seven metal layers
  - Package: 10/10/10μm W/S/T with 45° routing

- FFTx1 circuit contains 1.15 million cells in a 4mm square package
Our Benchmark 2D/2.5D Designs

- Multi-chip 2.5D FFT256 Designs
  - Duplicated the base design to implement FFTx2 and FFTx4 2.5D designs
Chip and package wires have significantly different dimensions
- Difficult for general chip extraction tools to perform extraction efficiently
- Irregular package routing increases extraction complexity
D2P coupling significantly increases with a closer D2P distance

<table>
<thead>
<tr>
<th>Top Layer</th>
<th>CCap (fF)</th>
<th>D2P distance (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>M7-R1</td>
<td>15.06</td>
<td>7.155</td>
</tr>
<tr>
<td>M6-R1</td>
<td>0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>M6-R1</td>
<td>4.330</td>
<td>3.319</td>
</tr>
</tbody>
</table>

E-field Shielding from RDL also affects D2P coupling

- Using one RDL layer is good enough because of ground planes
Ground Plane is NOT Enough

- What if we model the chip as a ground plane?
  - Compatible with most package tools (Allegro, Xpedition)

- With only ground plane, D2P coupling is significantly overestimated
  - Also, only lumped capacitance can be extracted

<table>
<thead>
<tr>
<th></th>
<th>CCap</th>
<th>Mtop</th>
<th>Via</th>
<th>rdl1</th>
<th>rdl2</th>
<th>rdl3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GND plane</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mtop</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.669</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>RDL1</td>
<td>4.669</td>
<td>0</td>
<td>0</td>
<td>0.782</td>
<td>41.44</td>
<td>0.145</td>
</tr>
<tr>
<td><strong>Real die</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mtop</td>
<td>60.92</td>
<td>2.793</td>
<td>1.019</td>
<td>0.075</td>
<td>0.055</td>
<td></td>
</tr>
<tr>
<td>RDL1</td>
<td>1.019</td>
<td>0.006</td>
<td>0.710</td>
<td>42.65</td>
<td>0.056</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ccap(pF)</th>
<th>Mtop-Mtop</th>
<th>Mtop-Via</th>
<th>Mtop-R1</th>
<th>R1-R1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Real die</strong></td>
<td>60.92</td>
<td>2.793</td>
<td>1.019</td>
<td>0.71</td>
</tr>
<tr>
<td><strong>Ground plane</strong></td>
<td>0</td>
<td>0</td>
<td>4.669</td>
<td>0.782</td>
</tr>
<tr>
<td><strong>Err%</strong></td>
<td>-100%</td>
<td>-100%</td>
<td>358%</td>
<td>10.2%</td>
</tr>
</tbody>
</table>
Holistic Extraction

- Traditionally, chip and package routings are extracted separately
  - D2P coupling is ignored completely
- Holistic extraction takes all metal layers from both chip and package into consideration
  - First, create a new metal stack configuration that includes all seven chip layers and three package layers.
  - Then, design a holistic die-package layout by merging the geometries from the die and the package.
  - Finally, run Calibre xACT on the holistic design to obtain all capacitance information
Holistic Extraction Results

- **D2P coupling capacitance extraction using holistic extraction**
  - Our GDS-level CAD tools can be used for complete D2P RCLM extraction

<table>
<thead>
<tr>
<th>Layer</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>rdl1</th>
<th>rdl2</th>
<th>rdl3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gcap</td>
<td>45856</td>
<td>24607</td>
<td>10502</td>
<td>6891.8</td>
<td>8429</td>
<td>2486.6</td>
<td>2316.7</td>
<td>17389</td>
<td>898.31</td>
<td>2833.8</td>
</tr>
<tr>
<td>Ccap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
  - M1 | 19093 | 21827 | 10267 | 1785 | 60 | 2.22 | 0.02 | 0.07 | 0 | 0 |
  - M2 | 21827 | 27895 | 53198 | 17424 | 377 | 20.68 | 0.24 | 3.85 | 8.55 | 5.23 |
  - M3 | 10267 | 53198 | 28434 | 88343 | 5005 | 185 | 0.87 | 3.31 | 4.4 | 3.23 |
  - M4 | 1785 | 17424 | 88343 | 33021 | 56288 | 5337 | 359 | 12.13 | 8.23 | 4.57 |
  - M5 | 60.23 | 377 | 5005 | 56288 | 18016 | 51222 | 375 | 31.08 | 12.46 | 6.96 |
  - M6 | 2.22 | 20.68 | 185 | 5337 | 51222 | 18708 | 16623 | 354 | 410 | 2.71 |
  - M7 | 0.02 | 0.24 | 0.87 | 359 | 375 | 16623 | 257 | 2519 | 47.05 | 0.23 |
  - rdl1 | 0.07 | 3.85 | 3.31 | 12.13 | 31.08 | 354 | 2519 | 134 | 31207 | 11.57 |
  - rdl2 | 0 | 8.55 | 4.4 | 8.23 | 12.46 | 410 | 47.05 | 31207 | 461 | 46256 |
  - rdl3 | 0 | 5.23 | 3.23 | 4.57 | 6.96 | 2.71 | 0.23 | 11.57 | 46256 | 231 |

Cap unit in fF
Coupling Cap Distribution

- Holistic extraction clearly captures all D2P coupling components
- With larger D2P overlapping area, D2P coupling increases significantly

![Graph showing coupling cap distribution with bars for M1-M7, R1-R3 indicating CCap and GCap with values in pF.]

- M1-M7 and R1-R3 bars show varying CCap and GCap values in pF.
- Right graph shows M7-R1 cap values for FFTx1, FFTx2, and FFTx4.
Holistic extraction is highly accurate but computationally expensive:
- it contains all layers from both the die and the package
- complicated structures on the package layers: hatch ground planes, pads, vias, 45° routing, circular shape routing
- requires in a significant increase in both runtime and memory

For example:
- two 10nm chip with 12 metal layers
- a three-layer package (20mm-by-20mm)
- 32-node (4 cores/node) computing grid
- more than 700GB of DRAM space
- more than three days
In-Context Extraction

- In-Context Extraction reduces both computational time and memory
  - In-Context extraction takes each chip separately
  - with a few interface layers from the neighboring component

- With the same example,
  - less than 150GB memory
  - less than 1.5 days

- Reuse most of the existing 2D LVS rule decks
  - Only some minor extensions are necessary for interface layers

- Eliminates the need for sharing intellectual properties across different design houses
  - Greatly simplifies the extraction procedure with heterogeneous designs
In-Context Design

- Using one RDL layer and cut each chip with D2P overlapping region
  - Preserves all D2P coupling and the compatibility with chip tools

![Image of die0 and die1]
In-context extraction still captures all D2P coupling elements and is highly accurate compared with holistic extraction.

- R1 coupling is smaller because of Fan Out region is not included.

<table>
<thead>
<tr>
<th>Method</th>
<th>Die0 + Die1 In-Context</th>
<th>Die0&amp;Die1 Holistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer</td>
<td>M7</td>
<td>M7</td>
</tr>
<tr>
<td>GCap</td>
<td>29.87</td>
<td>29.78</td>
</tr>
<tr>
<td>CCap</td>
<td>M7</td>
<td>M7</td>
</tr>
<tr>
<td>M7</td>
<td>121.8</td>
<td>121.9</td>
</tr>
<tr>
<td>Via</td>
<td>6.061</td>
<td>5.612</td>
</tr>
<tr>
<td>R1</td>
<td>2.202</td>
<td>2.187</td>
</tr>
</tbody>
</table>

The table above compares the coupling elements for in-context and holistic extraction methods. The values represent the coupling coefficients for different layers and connections, with M7, Via, and R1 being specific components of the coupling extraction process.
What about Inductance?

- **Halo-Ground Approximation**

  (a) Original layout
  (b) Halo ground layout

- **Bundle Creation**
Package-to-Chip RLM Estimation

- Less than 3s to compute more than 1000 package and chip wires
  - Accuracy verified against Fast Henry.

<table>
<thead>
<tr>
<th>#</th>
<th>R</th>
<th>TotL</th>
<th>TotM</th>
<th>R</th>
<th>ERR%</th>
<th>TotL</th>
<th>ERR%</th>
<th>TotM</th>
<th>ERR%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.52</td>
<td>34.3</td>
<td>94.8</td>
<td>0.52</td>
<td>-0.7%</td>
<td>35.3</td>
<td>2.8%</td>
<td>89.8</td>
<td>-5.3%</td>
</tr>
<tr>
<td>2</td>
<td>0.52</td>
<td>34.5</td>
<td>109.2</td>
<td>0.52</td>
<td>-0.6%</td>
<td>35.5</td>
<td>2.9%</td>
<td>104.6</td>
<td>-4.3%</td>
</tr>
<tr>
<td>3</td>
<td>0.42</td>
<td>27.5</td>
<td>106.3</td>
<td>0.41</td>
<td>-1.3%</td>
<td>26.8</td>
<td>-2.6%</td>
<td>96.4</td>
<td>-9.4%</td>
</tr>
<tr>
<td>4</td>
<td>0.52</td>
<td>34.8</td>
<td>121.1</td>
<td>0.51</td>
<td>-0.6%</td>
<td>35.8</td>
<td>3.0%</td>
<td>116.5</td>
<td>-3.8%</td>
</tr>
<tr>
<td>5</td>
<td>0.52</td>
<td>34.8</td>
<td>120.9</td>
<td>0.51</td>
<td>-0.6%</td>
<td>35.9</td>
<td>3.0%</td>
<td>116.3</td>
<td>-3.8%</td>
</tr>
</tbody>
</table>

0 1pH/μm

0 8pH/μm
Our tool identifies any noise hot-spot on chip and provides fast estimation of package-induced signal integrity issues.
Noise Analysis

- P2D inductive coupling noises are dominating
  - RCLM netlist is required for accurate noise estimation
Conclusion

- D2P coupling is heavily affected by E-field Shielding. Using ground plane approximation is not accurate enough for FOWLP.

- We propose holistic extraction with all metal layers and full geometries and demonstrate its maximum accuracy with high runtime and memory cost.

- We propose in-context extraction flows for reduced computational expenses and simplified flows for heterogeneous designs. In-context extraction is still highly accurate to capture D2P coupling.

- A comprehensive coupling extraction with RCLM netlist is needed for chip-package co-design, co-analysis, and co-optimization.