Jan 30, 2009

Research Challenges and Opportunities in 3D Integrated Circuits

Ankur Jain

ankur.jain@freescale.com, ankurjain@stanfordalumni.org
What is Three-dimensional (3D) Integration Technology?

► Main idea is very simple
  • Multiple layers of transistors / active devices
  • Connection between layers of active devices

► ITRS Roadmap Critical Path Challenge
  • Longest interconnect delay increasing with scaling
  • Limited opportunity to reduce delay by low-k
  • 3D integration can reduce interconnect length

► Manufacturing Options for 3D Circuits
  • Epitaxial growth of Si for monolithic integration
  • Die or wafer bonding of individually fabricated device-interconnect layers (Preferred)
Anatomy of a Bonded 3D Chip

3D Chip with three stacked dies

- Micro-bump
- Thru-Si via
- Cu filled TSV 52 um deep 15 um diameter

Bonding Orientations
1. Face-to-face: Metal interconnect sides of two strata bonded with metal micro-bumps
2. Back-to-face: Si side and metal interconnect side of two strata bonded with metal micro-bumps. Need through-Si vias (TSVs).
Inter-strata Connection Element 1: Micro-bump

3D Chip with three stacked dies

Metal Micro-bumps

Before bonding

After bonding (<300°C thermo-compression bond)

Typical foot-print 5 x 5 µm
20 – 60 µm pitch (compared to 170 – 200 µm C4 pitch)
Inter-strata Connection Element 2: Through-Si Via (TSV)

Side view of a Through-Si Via (TSV)

- Simplified Process steps:
  a) Etch hole from front side
  b) Deposit dielectric and barrier layer
  c) Fill hole with Cu
  d) Thin Si-substrate from back side to expose TSV

- For aspect ratio 10:1 or lower, typical foot-print 5 x 5 µm or larger with 50 µm deep via.

In addition to micro-bumps and TSVs, Inter-strata connection element can also consist of metal routing in an interposer die or metal routing from redistribution layer.
Micro-bump R and C characteristics

 Bonded micro-bump structure for 3D electro-static simulation

R and $C_{12}$ (inter micro-bump cap) vs. micro-bump size $pd$

Key Points

- Micro-bump size $pd$ $\downarrow$ $\Rightarrow$ $R \uparrow$, $C \downarrow$
- Max $C < 2fF$ for $pd$ of 35$\mu m$ x 35$\mu m$ with pitch=50$\mu m$
- For bump size = 5$\mu m$ x 5$\mu m$ with pitch=50$\mu m$, $R = 40$ m$\Omega$ & $C = 0.4$ fF

S. M. Alam et. al., ISQED, pp. 580-585, 2007
Through-Si Via (TSV) R and C Characteristics

▶ Closed form equations for a cylindrical Cu-filled TSV with height, \( h \), radius, \( r_{via} \), sidewall dielectric thickness, \( t \)

\[
R = \frac{\rho_{Cu} h}{\pi (r_{via} - 2 \cdot t)^2} \quad \text{... (1)}
\]

\[
C = \frac{2\pi \varepsilon_r \varepsilon_o h}{\ln\left((r_{via} + t) / r_{via}\right)} \quad \text{...... (2)}
\]

▶ Sidewall dielectric thickness, \( t \), is critical for \( C \)

Capacitance can be very high! Want large \( t \) for low capacitance

▶ \( h=50\mu m, r_{via} = 2.5\mu m, t=1\mu m \)

\( R = 43 \text{ m}\Omega, C = 0.04pF \)
Effect of Si Conductivity on TSV Capacitance

Equivalent circuit for Inter-TSV capacitance $C_{12}$

$C_{Si}$

$C_{d1}$  $R_{Si}$  $C_{d2}$  

TSV1  I  TSV2

Key Points

- As $\rho_{Si} \to \infty$, $R_{Si} \to$ open, $C_{12}$ saturates at lower value
- As $\rho_{Si} \to 0$, $R_{Si} \to$ short, $C_{12}$ saturates at higher value

- With typical $\rho_{Si} \sim 10$ $\Omega$-cm, Si acts like a resistor at low frequencies.

Inter-TSV capacitance, $C_{12}$, as a function of Si resistivity

Here, signal operating frequency, $f = 100$MHz, TSV dimension $s = 15$um, pitch $p = 30$um, dielectric thickness, $t = 0.2$um, and TSV height, $h = 50$um.
Effect of Frequency on TSV Capacitance and Inductance

Key Points

• Capacitance, C, dominates at low frequency
• Inductance, L, is much more complex to estimate due to dependence on return path
• Reported L range 0.3-0.9pH/um [5,7,8]
• If L ~ 10-50pH and C ~ 0.1-1pF, then resonant frequency:

  \[ \omega_0 = \frac{1}{\sqrt{LC}} \sim 100\text{GHz} \]

• Therefore, for low frequency, e.g. less than 1 GHz, operation, consider only R and C
Through-Si Via Scaling Trend

- TSV trades off device area on a Si substrate: TSV scalability is crucial

- TSV Scaling trend:
  - Density $\uparrow$ Footprint $\downarrow$ Height $\downarrow$
  - CMOS technology scaling

- RC product of a TSV remains the same (1$^{\text{st}}$ order) because of manufacturability constraints (constant AR and ratio of $t$, $r_{via}$)
Inter-strata Connection RC Delay and Scaling

Critical Dimension and RC parasitics of Inter-strata connections in 90nm CMOS technology

<table>
<thead>
<tr>
<th>Inter-strata Connection Element</th>
<th>Critical Dimensions in 90nm Node</th>
<th>Parasitic Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Micro-bump</td>
<td>Footprint: 5µm x 5 µm</td>
<td>R = 40 mΩ</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C = 0.4fF</td>
</tr>
<tr>
<td>Through-Si Via (TSV)</td>
<td>Footprint: 5 µm x 5 µm</td>
<td>R = 43 mΩ</td>
</tr>
<tr>
<td></td>
<td>Sidewall thickness, t: 1 µm</td>
<td>C = 40fF</td>
</tr>
<tr>
<td></td>
<td>Height, h: 50 µm</td>
<td></td>
</tr>
</tbody>
</table>

► Define inter-strata single-stage IO driver delay

► Compare with single-stage repeater delay in Global Interconnection

I. Savidis, et. al., VMIC, 2008
Delay Comparison with Global Interconnection

Global wire delay per mm length increases linearly with scaling in spite of increasing number of repeaters.

Single-stage global wire delay > Inter-strata IO driver delay with micro-bump + TSV ≅ average 5x t_{FO4}


I. Savidis, et. al., VMIC, 2008
3D Integration Reduces Wiring Latency and Power

Global wire delay per mm length with repeaters vs. 3D IO delay

Global wire power per mm length with repeaters vs. 3D IO power

Based on International Technology Roadmap for Semiconductors (ITRS) Roadmap and 800 MHz global wire frequency.
Thermal Management in Microelectronics

- The challenge of heat dissipation at die, package and system levels.
  - Localized, on-die heat generation (‘hotspots’).
  - Effective heat sinking technology.
  - System-level thermal-friendly design.

- Concept of thermal resistance (K/W).

- Application specific thermal management.
  - Liquid cooling/refrigeration for advanced computers.
  - Metal heat sink and air based cooling for PCs.
  - Passive cooling for ultra-low power products.

The current thermal management paradigm
Many novel thermal management strategies are currently being worked. Thermal management at the nanoscale is becoming increasingly important.

- Liquid/two-phase cooling
- Thin film thermoelectric cooling
- Advanced nanomaterials based cooling

... and several other technologies!
What makes 3D different?

► The thermal management challenge is exacerbated in 3D technology.
  • Reduced die footprint and increased power density.
  • Multiple heat sources.
  • Varying temperature requirements of multiple strata.
► The number of layers in stack will be limited by thermal considerations.
► Thermal performance of TSVs and micropad bonding layer needs to be understood.
► Need for early incorporation of thermal models in electrical CAD tools for 3D. **Thermally-aware** floorplanning, routing, TSV placement, etc.
Thermal considerations in 3D

- What is the maximum temperature and where?
- By how much will the heat sink thermal performance need to be improved to integrate two die with given power dissipations?
- Which parameters must be engineered to ensure minimal thermal impact of 3D integration?
- Is the traditional concept of junction-to-air thermal resistance sufficient for describing thermal characteristics of a multi-die stack?
One-dimensional Analytical Thermal Model

- Using the linearity of the energy equation, the solution for temperature distribution in a multi-die stack with multiple heat sources is found to be

\[
\theta_i = \frac{\sum_{k=1}^{i-1} Q_k (1-F_k) \left( 1 + \sum_{j=1}^{k} r_j \right) - F_k \sum_{j=k+1}^{i} r_j}{R_{hs}} + \sum_{k=i}^{N} Q_k (1-F_k) \left( 1 + \sum_{j=1}^{i} r_j \right)
\]

where \( \theta_i \) is the temperature rise above ambient, \( r_j \) is the thermal resistance of \( j^{th} \) layer, relative to heat sink resistance. \( \alpha \) is the non-dimensional package resistance.

- Eg. In a two-die stack, if each die dissipates 10W, the temperature rise is 43.8 K and 36.7 K. For 18W/2W power distribution, the temperature rise is 50.7 K and 35.3 K.

Jain, et al., IEEE Trans Comp Packagging Tech, accepted, 2009
Heat sinking requirements for 3D chips

- Figure shows the impact of 3D integration on heat sink performance requirements.

- Some difference between memory-on-logic and logic-on-memory.

Jain, et al., IEEE Trans Comp Packaging Tech, accepted, 2009
3D versus SoC

- 3D technology offers footprint and interconnection advantages over System-on-Chip (SoC). However, 3D also exacerbates the thermal management problem.

- The problem of cooling a 3D chip is neither much worse nor much easier than a 2D chip with the same footprint area.

- However, the 3D thermal resistance is higher than SoC for the same total area.

- 3D technology will require innovative thermal management solutions, both at die-level and package-level. Integration of thermal and electrical considerations will be required for resolution of trade-offs.

<table>
<thead>
<tr>
<th>Case</th>
<th>Junction-to-air thermal resistance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.8 mm by 7.8 mm SoC</td>
<td>12.8 K/W</td>
</tr>
<tr>
<td>7.8 mm by 7.8 mm 3D stack (same power; total area is double of the SoC; same footprint area)</td>
<td>13.5 K/W</td>
</tr>
<tr>
<td>5.5 mm by 5.5 mm 3D stack, (same power; total area is the same as SoC)</td>
<td>15.6 K/W</td>
</tr>
</tbody>
</table>

* based on max temperature rise and total power

Jain, et al., IEEE Trans Comp Packaging Tech, accepted, 2009
Thermal conductance of inter-die bond layer

- Thermal performance of the micropad bonding layer between the two die depends on geometry and thermal conductivity of the filler material and micropad metal.

- FE simulations show that this layer has fairly good thermal conductance.


Jain, et al., IEEE Trans Comp Packaging Tech, accepted, 2009
Due to higher thermal conductivity, it is possible to use a TSV as a ‘heat pipe’ and conduct heat away from areas with large local heat dissipation.

A 47% reduction in maximum temperature has been reported as a result of TSV insertion (Goplen & Sapatnekar, 2006; Cong & Zhang, 2005).

(a) Without Thermal Via Insertion  (b) With Thermal Via Insertion

Cong & Zhang, 2005
Thermal vs Electrical vs Mechanical Trade-offs

► TSV can aid in localized heat removal.

► However, it takes up valuable real estate on Silicon. Keep-out zone due to mechanical stress generation increases the real estate penalty.

► Typically, blocks with higher power dissipation will also need higher I/O, and thus more TSVs.

► Electrical-thermal-mechanical co-design is important for extracting the most thermal benefit out of TSVs.
Conclusions

► 3D technology offers both challenges and opportunities in thermal management
  
  • Increased power density, varying temperature specs of heterogeneous technologies.
  
  • Localized thermal management using TSVs.

► Thermal resistances internal to 3D technology are expected to be much smaller than external resistances like the package and heat sink. Thus, effective package-level thermal design will continue to be critical.

► Due to the strong electrical-thermal trade-off identified in 3D system design, early incorporation of thermal modeling in the CAD and layout process is essential.
Thus, die thinning or increasing micropad density may not offer much global thermal benefit. Overall thermal resistance continues to be dominated by that of the heat sink and/or package.
Thermal Characteristics with Reciprocal Design Symmetry

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Thermal resistance (K/W)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core-over-memory in (a) (80% power in Core)</td>
<td>Core1</td>
</tr>
<tr>
<td>Core-over-core in (b) (80% power in Core)</td>
<td>7.41</td>
</tr>
<tr>
<td>Core-over-memory in (a) (95% power in Core)</td>
<td>6.94</td>
</tr>
<tr>
<td>Core-over-core in (b) (95% power in Core)</td>
<td>7.80</td>
</tr>
</tbody>
</table>

** Kelvin per total power (W) in each die

 Thermal simulation set up in FloTherm for a Dual-core processor

Thermal advantages of **Reciprocal Design Symmetry** over core-over-core configuration:

- Peak temperature is minimized
- Inter-die and intra-die temperature gradient is minimized for less variations

Thermomechanical Reliability Test Results

► Demonstrated excellent reliability performance based on MSL, temperature cycling, temperature-humidity and high temperature storage tests.

► No intrinsic/systematic reliability issue. NUF & Bibenzyl dice show intrinsic reliability at MSL1 & MSL3 levels respectively.

<table>
<thead>
<tr>
<th>Test (NUF)</th>
<th>Duration</th>
<th>Parts Tested</th>
<th>15 µm Fails</th>
<th>25 µm Fails</th>
<th>35 µm Fails</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSL1</td>
<td></td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>85°C at 85% humidity</td>
<td>168 hours</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-40°C to 125°C cycling</td>
<td>2000 cycles</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>150°C high temp. storage</td>
<td>1000 hours</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-50°C to 165°C cycling</td>
<td>1000 cycles</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Micropads Lifetime– Electromigration, Kelvin structure

15um Pads

Probability Plot (Individual Fit) for TTF
Lognormal - 80% CI
Complete Data - ML Estimates

Temp
300
350

Table of Statistics
Loc Scale A D* F C
5.09633 0.940840 2.473 23 0
1.43159 0.963096 1.355 30 0

Temp
300
350

Table of Statistics
Loc Scale A D* F C
4.76895 2.13545 4.505

* Removed early fail

Ea = 1.06 ± 0.11 eV
Sigma = 0.45 ± 0.06
(Error based on 80% 2-sided CI)

Projected lifetime @105°C = 2990367 hours
(0.1% failure @105C with a 60% 1-sided CI)

25um Pads

Probability Plot (Individual Fit) for TTF
Lognormal - 80% CI
Censoring Column in Censor - ML Estimates

Temp
300
350

Table of Statistics
Loc Scale A D* F C
5.09633 0.940840 2.473 23 0
1.43159 0.963096 1.355 30 0

Temp
300
350

Table of Statistics
Loc Scale A D* F C
4.76895 2.13545 4.505

* Removed early fail

Ea = 1.00 ± 0.25 eV
Sigma = 1.03 ± 0.16
(Error based on 80% 2-sided CI)

Projected lifetime @105°C = 126249 hours
(0.1% failure @105C with a 60% 1-sided CI)

Z. Huang, et al., ECTC, 2008