Adding a New Dimension to Physical Design

Sachin Sapatnekar
University of Minnesota
Outline

• What is 3D about?

• Why 3D?

• 3D-specific challenges

• 3D analysis and optimization
Planning a city: Land usage

[Somewhere in the American midwest; pop. density typically about 20 persons/km²]

[Minneapolis, p.d. = 2,700/km²]  [SF= 6,688/km²]  [New York=10,600/km²]
Types of 3D circuits

- **PCB stacking**
- **Memory - vertical TFTs**
- **Wafer stacking**

**Example application**
- Antenna Layer
- LNA / Mixer
- Down-conversion layer: IF, ADC, Digital Baseband
- Digital processing
- Isolation plane
- Back-Metal

[Matrix Semiconductor]

[www.irvine-sensors.com]
Example of a commercial application

STMicroelectronics CMOS camera

Febr. 2008 - STMicroelectronics announces that its single-chip camera sensors are now available in ST’s TSV (Through Silicon Via) wafer-level package technology. Unit pricing is in the $2 range, depending on the production period and quantities. ST’s VD6725 single-chip camera sensor is available in two package options, as a COB (Chip On Board) die or in the TSV wafer-level package. The sensor fits in phone camera modules smaller than 6 x 6 x 3.8 mm thanks to its 1.75-micron pixel design and ST’s advanced sensor architecture.
Example 3D processes

[Koyanagi, Tohoku U./Zycube]

[Hedler, Qimonda]

[IBM]
Through-silicon vias (TSVs)

Keep-out distance

[Tezzarron]

[Nowak, Qualcomm]
Schematic of a 3D IC

Adapted from [Das et al., ISVLSI, 2003] by B. Goplen
Outline

• What is 3D about?

• Why 3D?

• 3D-specific challenges

• 3D analysis and optimization
Another “dimension” to scaling

3D provides an alternative avenue towards increasing system sizes

Orthogonal to device scaling
3D Interconnects

- Reduced wire lengths

- Theoretically
  - For an $L \times L$ 2D chip, max wire length reduces from $2L$ to $\frac{2L}{\sqrt{m}}$

![3D Global Net Distributions](image)
Why are shorter wires good?

- Sequential critical length ("cycle reach") trends

- Critical interbuffer length also shrinking (i.e., buffer count increasing)
Other benefits

- Improved isolation in 3D
  - Critical for analog/RF ckts
  - Lower digital/mixed-signal noise
  - Shielding is possible either using metal layers, or by leveraging bonding material

- Heterogeneous integration
  - Different layers can be made of different materials
  - Can integrate, for example
    - CMOS
    - GaAs
    - Optical elements (VCSELs)
    - MEMS/NEMS
    - Exotic cooling technologies (micropumps, piezoelectric devices, microrefrigerators)

[Das et al., ISPD04]
Outline

• What is 3D about?

• Why 3D?

• 3D-specific challenges

• 3D analysis and optimization
Geometrical challenges

Adapted from [Das et al., ISVLSI, 2003] by B. Goplen
Thermal challenges

• Each layer generates heat
• Heat sink at the end(s)

• Simple analysis
  – \( \text{Power(3D)}/\text{Power(2D)} = m \)
    • \( m = \) # layers
  – Let \( R_{\text{sink}} \) = thermal resistance of heat sink
  – \( T = \text{Power} \times R_{\text{sink}} \)
    • \( m \) times worse for 3D!

• And this does not account for
  – Increased effective \( R_{\text{sink}} \)
  – Leakage power effects, T-leakage feedback

• Thermal bottleneck: a major problem for 3D
Thermal impact on circuit performance

- **Gate delays** change with T
  - Mobility goes down
    \[ \mu(T) = \mu(T_0) \left( \frac{T}{T_0} \right)^{-m} \]
  - \( V_{th} \) goes down
    \[ V_{th}(T) = V_{th}(T_0) - \kappa(T - T_0) \]
  - Which effect wins?
  - Positive, negative, mixed T dependency

- **Wire delays** change with T
- **Leakage** increases with T
- **Reliability** degrades with T
  - NBTI, electromigration

- Can use better heat sinks, but…

*The same circuit at various process corners*

*Heat sink cost vs. Power*
Power delivery challenges

• Each layer draws current from the power grid
  • Power pins at the extreme end tier(s)

• Simple analysis
  – Current(3D)/Current(2D) = $m$
    • $m = \# \text{ layers}$
  – Let $R_{grid} = \text{resistance of power grid}$
  – $V_{\text{drop}} = \text{Current} \times R_{grid}$
    • $m$ times worse for 3D!

• And this does not account for
  – Increased effective $R_{grid}$
  – Leakage power effects, increased current due to T-leakage feedback

• Power bottleneck: a major problem for 3D
Power supply integrity in 3D

- Greater challenge in 3D due to via resistance, limited number of supply pins

**Current per power pin (2D) – ITRS**

![Diagram](image)

*Current = 2I  
Single-story logic  
Current = I  
Multi-story logic*

[Zhan, ICCAD07]
Yield/test challenges

- Yield due to spot defects reduces exponentially with area
  - Smaller areas imply better yield
  - Stack together smaller die; yield improves!
  - (Note that stacking wafers together does not help!)

- Problem
  - Need to have known-good die (KGD)
  - Must test die prior to 3D assembly

- Testing thinned die is hard: mechanically too weak for probe pressure!
- Can test die prior to thinning – but then, connections to other layers are untested!

[Mak, Intel]
Outline

• What is 3D about?
• Why 3D?
• 3D-specific challenges
• 3D analysis and optimization
Thermal analysis

• Heat generation
  – Switching gates/blocks act as heat sources
  – Time constants for heat of the order of ms or more

• Thermal equation: partial differential equation

\[ K_x \frac{\partial^2 T}{\partial x^2} + K_y \frac{\partial^2 T}{\partial y^2} + K_z \frac{\partial^2 T}{\partial z^2} + Q(x, y, z) = 0 \]

• Boundary conditions corresponding to the ambient, heat sink, etc.

• Self-consistency
  – Power = f(T)
  – T = g(Power)
Thermal solution techniques

- Numerical: solve large, sparse systems of linear equations
  - Finite difference method: thermal – electrical equivalence
    - System structure is similar to power grids (good!)
    - Current sources ↔ power, voltage ↔ temperature
    - Finite element method
- Semi-analytical
  - Green functions (fast, appropriate for early analysis)
Thermal optimization

- Minimize power usage
- Rearrange heat sources
- Improved thermal conduits
- Improved heat sinking
3D floorplanning

Model the 3D chips and blocks

Thermal-aware lateral spreading

Global optimization in continuous 3D space

Layer assignment

3D force-directed optimization

Layer assignment

Final legalization

[Zhou, ICCAD07]
3D placement

- Incorporate thermal issues
- Force-directed vs. partitioning methods
Interlayer via count vs. wirelength (ibm01)
Thermal vias

- Thermal vias
  - Electrically isolated vias
  - Used for heat conduction

- Thermal via regions
  - Contains thermal vias
  - Predictable obstacle for routing
  - Variable density of thermal vias
Temperature profile

Before Thermal Via Placement

After Thermal Via Placement
Thermal via insertion

Temperature Profile

Thermal Via Regions
3D routing with integrated thermal via insertion

- Build good **heat conduction path** through dielectric:
  - **Thermal vias**: interlayers vias dedicated to thermal conduction.
  - **Thermal wires**: metal wires improves lateral heat conduction.
  - **Thermal vias + thermal wires** → a thermal conduction network.

- Thermal wires compete with lateral signal wire routing.
- Thermal vias: large, can block lateral signal routing capacity.

[Zhang, ASPDAC06]
Active cooling techniques

[Diagram showing Si Die, Polymer cover, Si microchannel heat sink, TWV, Fluidic I/O, Electrical I/O, Optical I/O, "Trimodal I/O"]

[Bakir, GaTech - CICC 07]
Microfluidic cooling

Trimodal I/Os

TSV-E

TSV-F

Fluidic channel

Cu wire

Optical waveguide

Temperature rise on heaters (°C)

Localized power density (W/cm²)

Flow rate = 34 ml/min
Flow rate = 78 ml/min
Flow rate = 104 ml/min
Flow rate = 125 ml/min

Area: 8 mm²

Flow rate

[Area: 8 mm²]
3D and multicore systems

On-die Mesh Interconnect

Processor Tile

Memory Bus

Signals and power from package, through memory, to the processor tile

<table>
<thead>
<tr>
<th>TSV Pitch</th>
<th>190μm</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRAM die size</td>
<td>275mm²</td>
</tr>
<tr>
<td>SRAM size</td>
<td>256KB per tile, 20MB total</td>
</tr>
<tr>
<td>SRAM Power</td>
<td>7W SRAM + 2.2W IO</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>12GB/sec/tile, ~1TB/sec total</td>
</tr>
</tbody>
</table>

[Karnik, Intel]

NoCs

Audio NIC
VIDEO NIC
CPU NIC
L2 $ NIC
CPU NIC
CPU NIC

Processing Element (PE)

b-bit Links

Router

3D bus/NoC hybrid

4x4x4 3D NoC-Bus Hybrid NoC

6x6 Crossbar

[Xie, Penn State]
3D NoCs

- Need to build custom NoCs for 3D architectures
- Floorplanning + NoC design
- 3D-specific challenges
  - Technology constraints, like TSV#
  - Tier assignment
  - Placement of switches
  - Accurate power and delay modeling

[Zhou, ASPDAC10]
Conclusion

• Numerous challenging problems in 3D IC design
• Significant research already in floorplanning, placement, routing
• New challenges in architectural-level issues, NoCs, power delivery, test
Thank You!
Any Questions?
You!
Thank