3D Heterogeneous Technologies for [Memory-Processor] & [CMOS-Sensor] Stacking

Muhannad S. Bakir
School of ECE, Georgia Tech
Nanoelectronics Research Center (NRC)

muhannad.bakir@mirc.gatech.edu

Future Directions in Packaging (FDIP) 2010
Turn of the Century Marked Many New Paradigms …

On-chip wires dominate:
- Latency & energy dissipation exceed transistor
- More masking levels

'Simple' scaling has ended:
- Lithography +strain + high-k
  (+ V_{dd}/V_t scaling slowed)
- non-planar CMOS next???

Historic frequency scaling ended
- design complexity
- energy efficiency
- but, no “free lunch”
Many-Core Processor Emergence

Aggregate off-chip bandwidth:
- Today: ~0.8 Tbps
- Soon: Several Tbps

Few challenges:
- Interconnect quality & density
- Latency (inches of wire)
- Power:
  - ~15-20% μP power used for signal I/Os
Some Challenges in Off-Chip Signaling

**Freq. Dependent Losses**

![Graph showing frequency dependent losses.](image)

- Loss inc by 5.5%-49.5% at 5 GHz

**Copper Surface Roughness**

- Card A- Ground Layer

**Not Enough Signal Pins**


**A Whole Lot of Discontinuities**

- Capacitance
- Impedance changes and discontinuities
- Inductance
- Signal stub and reflections
- Cross talk
- Dielectric loss, material inconsistencies

*SiliconPipe website*

---

B. Casper et al. CICC 2007
Energy Cost for Off-Chip Communication

<table>
<thead>
<tr>
<th>Operation</th>
<th>Energy (pJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>64b Floating FMA (2 ops)</td>
<td>100</td>
</tr>
<tr>
<td>64b Integer Add</td>
<td>1</td>
</tr>
<tr>
<td>Write 64b DFF</td>
<td>0.5</td>
</tr>
<tr>
<td>Read 64b Register (64 x 32 bank)</td>
<td>3.5</td>
</tr>
<tr>
<td>Read 64b RAM (64 x 2K)</td>
<td>25</td>
</tr>
<tr>
<td>Read tags (24 x 2K)</td>
<td>8</td>
</tr>
<tr>
<td>Move 64b 1mm</td>
<td>6</td>
</tr>
<tr>
<td>Move 64b 20mm</td>
<td>120</td>
</tr>
<tr>
<td>Move 64b off chip</td>
<td>256</td>
</tr>
<tr>
<td>Read 64b from DRAM</td>
<td>2000</td>
</tr>
</tbody>
</table>

(Bill Dally, Advanced Computing Symposium, September 16, 2009)

- Lots of energy goes into communication
- Need new approach to performing chip-chip signaling
3D Stacking – with Air-Cooled Heat Sink

(+) interconnect length (1,000x length reduction)
(+) interconnect energy (10-40x lower)
(+ /--) system footprint
(-) power delivery for processor
(-) memory density
(-) number of memory chips
3D Stacking – with Interlayer Liquid Cooling

(+) interconnect length (1,000x length reduction)
(+) interconnect energy (10-40x lower)
(+) interconnect density (100x easily)
(++) system footprint
(++) power delivery for processor
(++) memory density
(++) number of memory chips
Impact of Novel Cooling on Power Dissipation

\[ P = \frac{1}{R_{\text{thermal}}} (T - T_{\text{ambient}}) \]

\[ P = aC_{\text{total}} \left( V_{dd}(T) \right)^2 f + N_{\text{gates}} V_{dd}(T) I_{\text{leak}} e^{-\frac{V_t(T)+\Delta V_t}{nkT/q}} \]

<table>
<thead>
<tr>
<th></th>
<th>Freq.</th>
<th>Power</th>
<th>Temp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air cooling: ~0.6 °C/W</td>
<td>3 GHz</td>
<td>102 W</td>
<td>88 °C</td>
</tr>
<tr>
<td>Advanced Cooling: ~0.25 °C/W</td>
<td>3 GHz</td>
<td>83 W</td>
<td>47 °C</td>
</tr>
</tbody>
</table>

- Thermal ‘ancillary technologies’ are critical to minimizing power dissipation and increasing reliability and performance

D. Sekar et al., IITC 2008
B. Dang et al., IEEE TAP 2010
Sekar et al, IITC 2008
Interlayer Microfluidic Cooling Approach for 3D ICs

- Direct cooling of chip back side eliminates thermal interface resistance
- Cooling on each stratum to extract >100W/cm²
- Microscale fluidic interconnection between strata

C. King et al., ECTC 2010
J. Zaveri et al., IMAPS 2009
Experimental Results

Microfluidic Channels

Electrical TSVs

(a) (b)

X Ray Image

Optical Microscope Image

J. Zaveri et al, IMAPS 2009
µScale Plumbing

Fluidic & Fluidic &

Fluidic &

Elec. I/O

Liquid-cooled Board
Air-gap C4 based Electro-fluidic I/Os

- Air-gap C4 I/Os – Solder based electro-fluidic interconnections
- Similar to C4 based I/Os except the solder is overplated on the mold which domes after reflow
- All advantages of conventional C4 I/Os
- In addition, it enables ability to use no-flow under fill during flip chip bonding

C. King et al., ECTC 2010
Assembly of C4 Electrical and Electro-fluidic I/Os

SEM  X-ray

• X-ray image of C4 Fluidic I/Os after assembly
→ 100 μm diameter fluidic TSVs
→ 45 μm tall C4 fluidic I/Os
→ 47 μm tall electrical I/Os

C. King et al., ECTC 2010
Fluidic Testing

- Flow rate measured up to 100ml/min

C. King et al., ECTC 2010
J. Zaveri et al, IMAPS 2009
Assuming Only the Top Most Die is Switching

- Technology node: 45nm
- Package inductance: 0.5nH
- Decap: 20% chip surface
- 100 A/cm² per chip

![Diagram of stacked dies and package]

- When the number of chips is increased, the noise is suppressed because non-switching dice provide additional decap
- When too many dice are stacked, on-chip decap can’t compensate the longer inductive TSVs.

One Layer Switching is Too Idealistic

- To maximize gain from 3D, blocks with most communication must be vertically interconnected.

- Therefore, we must consider the worst case when all the layers are switching.
• When all the dice are switching, the noise condition in the 3D stack is unacceptable compared to single chip case, especially for the top most die.

• We need to find ways to suppress noise!
3D Problem Needs a 3D Solution: Use of a ‘Decap’ Die

|V_{noise}|=400 mV

“Decap” die: 100% decap.

|V_{noise}|=312 mV, 22% reduction

|V_{noise}|=256 mV, 36% reduction

→ functions as a local high-frequency energy storage

G. Huang et al., EPEP 2007
But, there is more to the story …

3D Stacking of Electronics and MEMS/sensors
MEMS Market

“The number of different MEMS devices is large and steadily growing”

2004 Total $12 billion

2009 Total $25 billion

Table 3: Forecasted growth (in $bn) of MEMS devices

<table>
<thead>
<tr>
<th>Application</th>
<th>2005</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure sensors</td>
<td>3.0</td>
<td>6.0</td>
</tr>
<tr>
<td>In vitro diagnostics</td>
<td>0.01</td>
<td>5.0</td>
</tr>
<tr>
<td>Read/write heads</td>
<td>2.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Ink jet print heads</td>
<td>2.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Optical displays</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Gyroscopes</td>
<td>0.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Lab-on-a-chip</td>
<td>0.01</td>
<td>2.0</td>
</tr>
<tr>
<td>Drug delivery systems</td>
<td>0.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Inertial sensors</td>
<td>0.2</td>
<td>1.5</td>
</tr>
<tr>
<td>Chemical sensors</td>
<td>0.1</td>
<td>1.0</td>
</tr>
<tr>
<td>Optical switches</td>
<td>0.1</td>
<td>1.0</td>
</tr>
<tr>
<td>RF devices</td>
<td>0.1</td>
<td>1.0</td>
</tr>
<tr>
<td>Microspectrometers</td>
<td>0.02</td>
<td>0.4</td>
</tr>
</tbody>
</table>

*T. Marinis, Strain 2009
MEMS/Sensor and Need for Electronics

• MEMS/Sensors need electronics
  – Signal conditioning, amplification, analysis, device actuation, etc.

• Challenges for MEMS/electronics monolithic integration:
  – Most of the state-of-the-art foundry wary of preprocessed wafers
  – MEMS last approach gives limited window to MEMS designers
    – Limited processes, materials, and devices
  – Each monolithic process is unique
  – Increases the development time as well as NRE cost
  – Supply voltage
  – However, provides small electrical parasitics
Package Based / Hybrid Integration

- MEMS and CMOS can be manufactured independently
  - Shorter Time-to-Market
  - Less complex process
  - Lower Non-Recurring Expense
- Low performance due to 2D interconnects
  - Signals need to go through several millimeter of wire and package wires

![Diagram of MEMS and CMOS integration](image)

Analog Devices’ ADXL345 Package X-Ray Image from MEMS the Word
Heterogeneous 3D Integration

- Independent fabrication of CMOS and MEMS
- Performance benefits of 3D integration

There is a need for new interconnect technologies
Mechanically Flexible Interconnects

- Stress Isolation
- Assembly on non-planar surface
- Potential temporary interconnections for disposable sensors

H S Yang et al ECTC 2010
Disposable Sensors

- If temporary interconnections are possible...
- Cleaning the sensor is expensive if not impossible
- Cost per test can be reduced by reusing CMOS IC

H S Yang et al. IITC 2010
R. Ravindran et al. ECTC 2010
Compliance Measurements

Hysistron Triboindenter

Indenter Tip

Displacement vs. Force Graph

MFIs
Compliance vs. Thickness

H S Yang et al ECTC 2010
Bio Detection in General

• To detect a particular antigen, its complimentary antibody is first bound to the sensor
  – Antigen refers to cancer markers and antibody to proteins which specifically bind the cancer markers
• Then sample is introduced
  – The antibody specifically binds the antigen if present
  – The rest of the sample is washed away
Nanowire Based Sensors

• Nanowire (NW) sensors detect the charge induced due to the presence of charged proteins bound to their surface

• Surface charge leads to either an accumulation or depletion of carriers

• Analogous to surface potential in a FET controlling depletion depth and the onset of inversion

Figure retrieved from I. Kimukin et al., Nanotechnology 17, S240 (2006).
Silicon Nanowire Biosensor - BioFET

- A top-down fabricated charge-based sensor
- Label-free detection possible
- Can be used to detect cancer cells

Figure retrieved from P. Nair et al., IEEE Trans. Elec. Devices 54, 3400 (2007).
BioFET - TSV Integration

- Integration with a high temperature (900C) SiNW process demonstrated
- SiNW can be **fabricated prior to TSV** due to CMP-free planarization

H S Yang et al IITC 2010
R. Ravindran *et al.* ECTC 2010
Disposable Sensors

Contaminated Sensor

New Sensor Die

CMOS IC
Si Nanowire Sensor: **Samples from Real Patients**

- Shows detection of cancerous epithelial ovarian cells
- HEY -> Cancerous Cells
- IOSE -> Healthy Cells

R. Ravindran et al. ECTC 2010

*Collaboration w/ John McDonald, Biology GT & Atlanta Northside Hospital*
Conclusion

1) Innovation in silicon technology without REVOLUTIONARY innovation in silicon ancillary technologies will yield progressively “performance crippled” electronic systems.

2) 3D stacking & novel silicon ancillary technologies are key to:
   - Advancing computing systems
   - Enabling heterogeneous integration of electronics & MEMS/sensors