Cost and Performance Effective Solution for 3D ASIC and Memory Integration

Li Li
Cisco Systems, Inc.
Introduction – Insatiable Need for Bandwidth

- By 2018, there will be 10 billion mobile devices and connections

Internet traffic growth by device - 2012-2017 CAGR
Source: CISCO VNI 2013

DRAM cell will stop scaling at 1znm
Internet of Everything – Next Big Thing

...By connecting People, Process, Data and Things

People
Connecting people in more relevant, valuable ways

Process
Delivering the right information to the right person (or machine) at the right time

Data
Leveraging data into more useful information for decision-making

Things
Physical devices and objects connected to the Internet and each other for intelligent decision-making; often called Internet of Things (IoT)

IoE
Network Applications

Network Processor (NPU)
- Traffic Manager
- Packet Processing
- Switch Buffer

"The Sea of Memory"
System Implications

• **Performance**
  – Aggregated bandwidth of the NPU package needs to scale with system throughput requirements
  – “Sea of Memory”

• **Pins**
  – Package size and terminal pitch have been roughly flat
  – Number of I/O pins are limited
  – DDR memory data rate is not keeping with application needs

• **Power**
  – Total system-level power budget has been relatively flat
  – Customers expect 2X improvement in performance/watt

*3D Integration can address system level challenges with Performance, Power and Pins*

“Sea of Memory” → “Stacks of Memory”
<table>
<thead>
<tr>
<th>FCAMP</th>
<th>2.5D MCM-TSV</th>
<th>3D SiP-TSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Single bare ASIC die</td>
<td>• Multiple bare dice on one side of the interposer with TSV</td>
<td>Multiple bare dice on both sides of the interposer with TSV</td>
</tr>
<tr>
<td>• Packaged memory devices</td>
<td>• Higher wiring density</td>
<td>Higher IC integration</td>
</tr>
<tr>
<td>• Large package substrate size to include many components</td>
<td>• Optimized thermal performance</td>
<td>Higher wiring density</td>
</tr>
<tr>
<td>• High thermal performance</td>
<td></td>
<td>Short, direct interconnect</td>
</tr>
<tr>
<td>Wiring density limited by the build-up technology</td>
<td>Interposer size is often limited to 26 mm x 32 mm</td>
<td>May require trade-off between performance and power</td>
</tr>
</tbody>
</table>

- In production
- In development
- In development

 ![Diagram of 2.5D / 3D System Integration](image)
Si-Interposer Manufacturing & Supply Flows

“Foundry Process”

Foundry → TSV & FEOL → Front Side uBump → Wafer Thinning → Back Side Bump → Debond → Assy / Ship

Ship to OSAT

“MEOL Process” at OSAT

Foundry → TSV & FEOL → Front Side uBump → Wafer Thinning → Back Side Bump → Debond → Assy

“Substrate Process”

Subs Supplier → TSV & FEOL → Front Side uBump → Wafer Thinning → Back Side Bump → Debond → Ship
Emerging High Performance Memory

- Wide I/O
- Low Power
- Serial I/O
- Higher Power

Diagram showing the total bandwidth (GB/s) versus I/O data rate per pin (Gb/s) for different memory technologies and their performance over time from 2012 to 2015.
Assembly and Packaging for 3D SiP Modules

“Cost and Performance Effective Silicon Interposer and Vertical Interconnect for 3D ASIC and Memory Integration”, Session 31: PoP, SiP, and Die Stacking, Friday, May 30, 2014, 1:30 PM - 5:10 PM
Summary

• The insatiable needs for bandwidth has huge implications on the next gen network systems.

• 2.5D / 3D ASIC and Memory integration is promising and has high potentials to address some of the challenges seen.

• To enable 2.5D / 3D ASIC and Memory integration, a completed supply chain ecosystem / platform is needed.

• In addition to the development in memory components and interposers, driving innovations across the boundaries of suppliers should be encouraged.