3D TECHNOLOGIES: SOME PERSPECTIVES FOR MEMORY INTERCONNECT AND CONTROLLER

CODES+ISSS: Special session on memory controllers

Taipei, October 10th 2011

Denis Dutoit, Fabien Clermidy, Pascal Vivet

{denis.dutoit@cea.fr}

MINATEC CAMPUS

3D-IC design: will dream become reality ?

- Example of an heterogeneous 3D-IC stack:

Stackable Processor

Standard DRAM
Parallel & low speed IF

DRAM

Proc

Proc

Standard Non Volatile Memory
Serial high speed IF

RF/Analog

NVM

Active interposer (mature technology): Services & communication infrastructure ;
IO with ESD; Test; Supply distribution; Power management

- This presentation focuses on vertical interconnect between complex MPSoC and memory

© CEA. All rights reserved

Outline

- Introduction
- 3D TSV technology choices
  - 3D-IC toolbox
  - Design perspectives
- Improving memory bandwidth
  - The memory link bottleneck
  - Overview of memory interfaces: bandwidth, power consumption, package integration
- Wide IO interface integration: Mag3D circuit
  - Wide IO technology
  - Mag3D circuit: architecture, Wide IO DRAM integration, design implementation challenges and current results
- Conclusion & perspectives

© CEA. All rights reserved

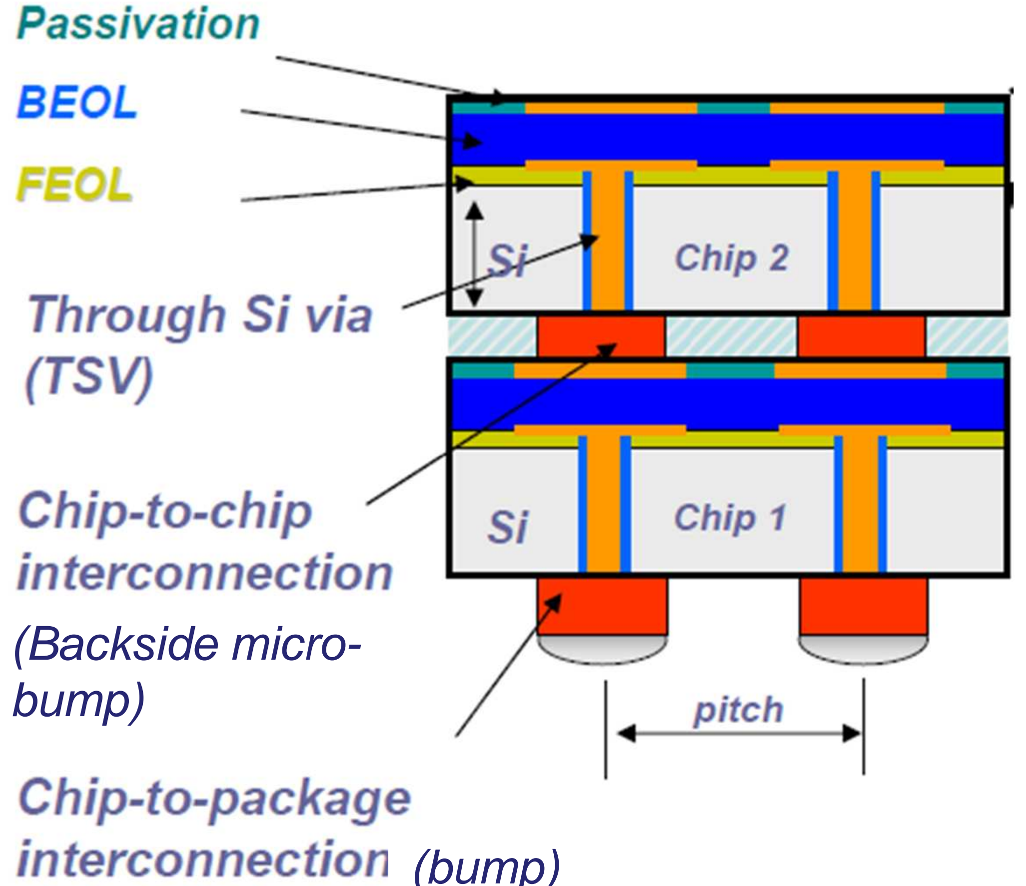
Outline

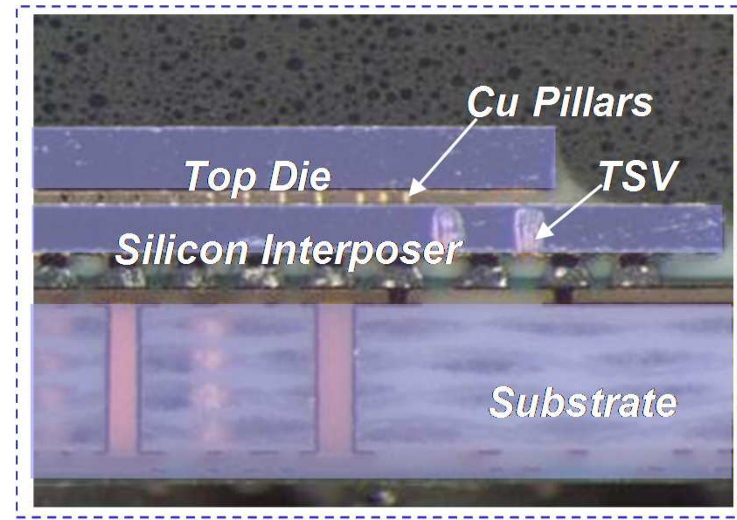
- Introduction
- 3D TSV technology choices
  - 3D-IC toolbox
  - Design perspectives
- Improving memory bandwidth
  - The memory link bottleneck
  - Overview of memory interfaces: bandwidth, power consumption, package integration
- Wide IO interface integration: Mag3D circuit
  - Wide IO technology
  - Mag3D circuit: architecture, Wide IO DRAM integration, design implementation challenges and current results
- Conclusion & perspectives

© CEA. All rights reserved

3D IC: Some definitions

- 3D-IC chip stacking: Adding a new vertical interconnect chain in CMOS : Through Silicon Via + backside bump





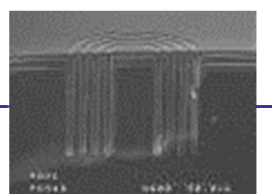
BEOL: Back-End-Of-Line
FEOL: Front-End-Of-Line

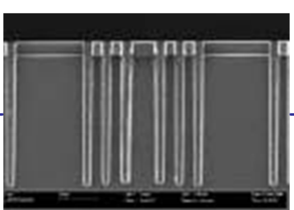
© CEA. All rights reserved

3D IC interconnect: Through Silicon Via

Via First TSV (Polysilicon filled)

Processed before CMOS front-end steps
Pitch: ~10µm
Density: 10000 TSV/mm²

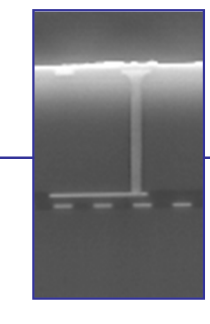




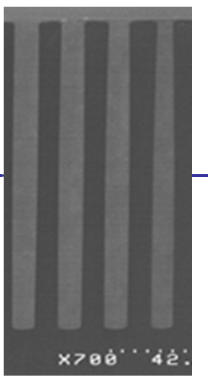
Trench AR 20,
5x100µm

Via Middle TSV (Copper filled)

Processed after CMOS front-end steps
Pitch: 40µm to 50µm
Density: 500 TSV/mm²



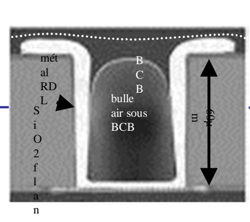
AR 7 , 2 x 15µm



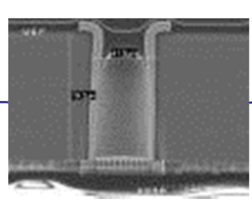
AR 10,
10x100µm

Via Last TSV (Copper liner)

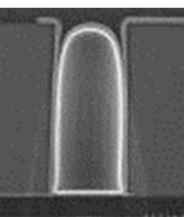
Processed after metallization
Pitch: ~100µm
Density: 100 TSV/mm²



AR 1
80x80µm



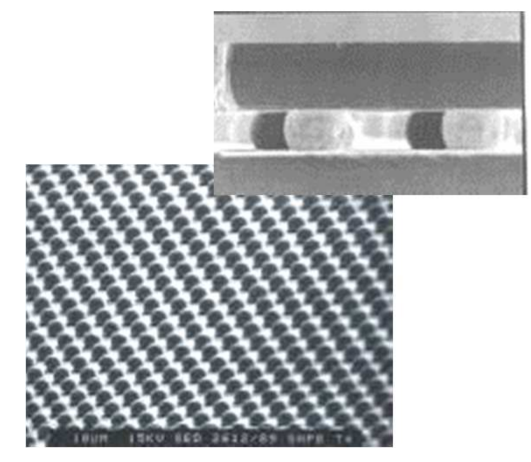
AR 2,
60x120µm



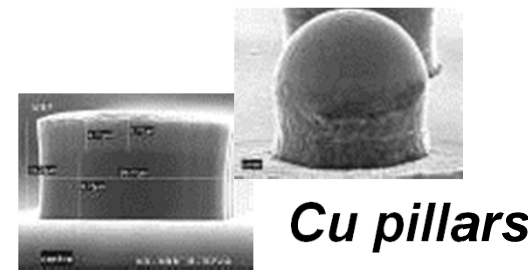
AR 3,
40x120µm

© CEA. All rights reserved

3D IC interconnect: inter-chip connections

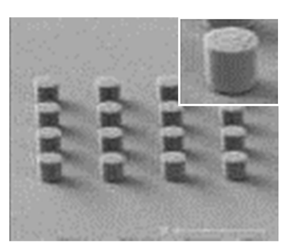


Classic Flip chip (Ball or stud bump)



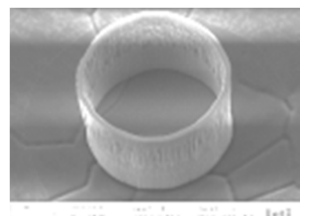
Cu pillars

SnAg Solder



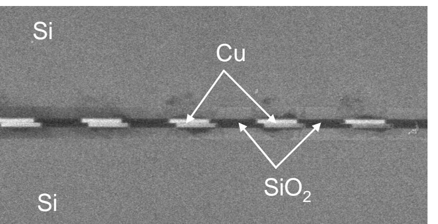
Solder-free µinserts

Nickel Microinsertion



µtubes

Microinsertion



Cu-Cu Direct bonding

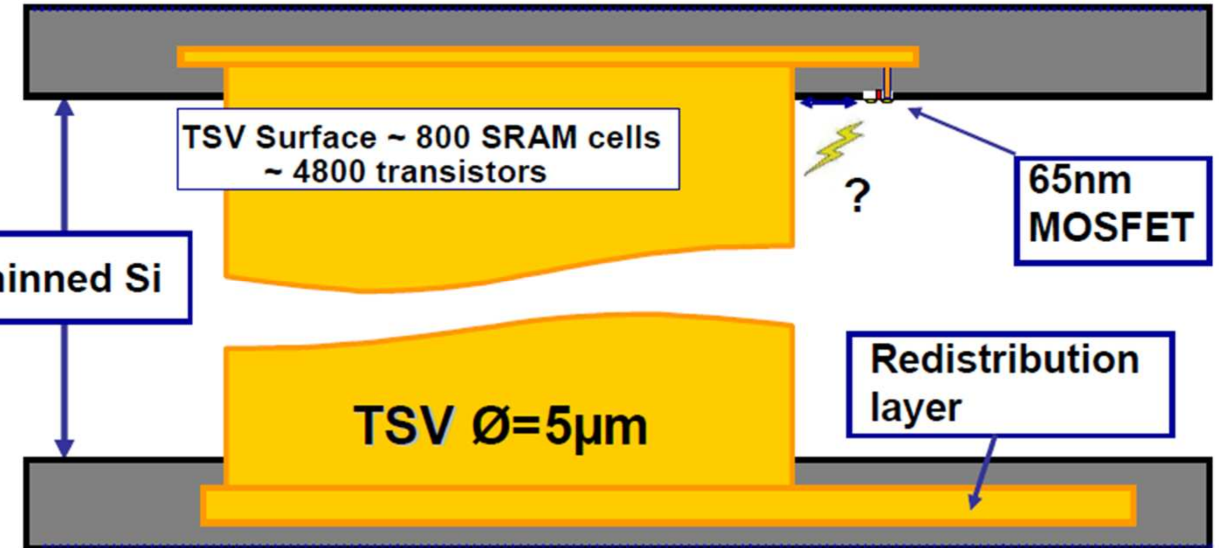
Need very flat surface
Cleaning

> 100 µm — *100-30 µm range* — *30-10 µm range* — *Down to 5 µm*

TSV last — ***TSV middle*** — ***TSV first*** — ***Pitch***

© CEA. All rights reserved

Current TSV technology and design perspectives



1. What can we expect ?
   - 10 µm diameter TSV is challenging
   - Must add guard interval
   - Die thickness is an issue
2. 3D design is not yet mature:
   - No accurate CAD tools & models for 3D (CTS, extraction, Thermal, Power …)
   - No accurate test model for TSV

- Resulting 3D partitioning ?
  - Fine grain partitioning of an operator in two dies using TSV is not a good option

  ⇒ **Coarse grain to medium grain partitioning**

⇒ **Proposal ? 3D Memory (Wide IO)**

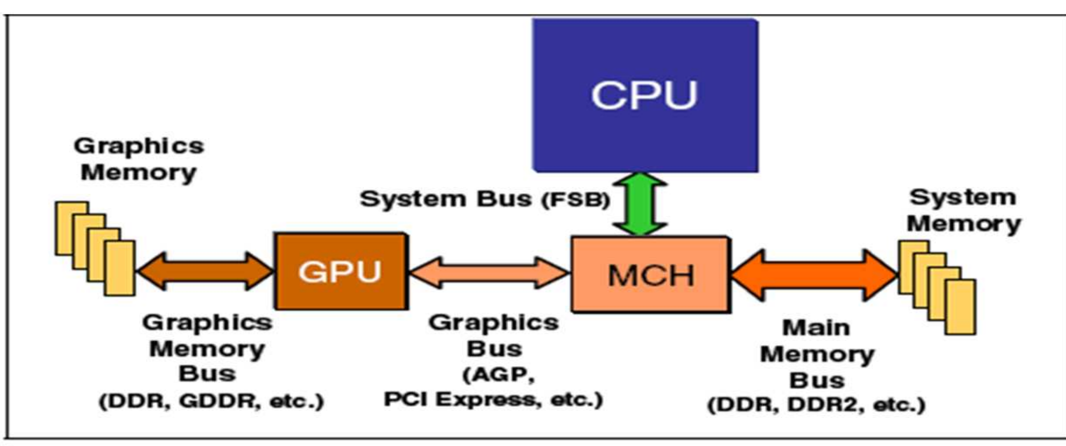
© CEA. All rights reserved

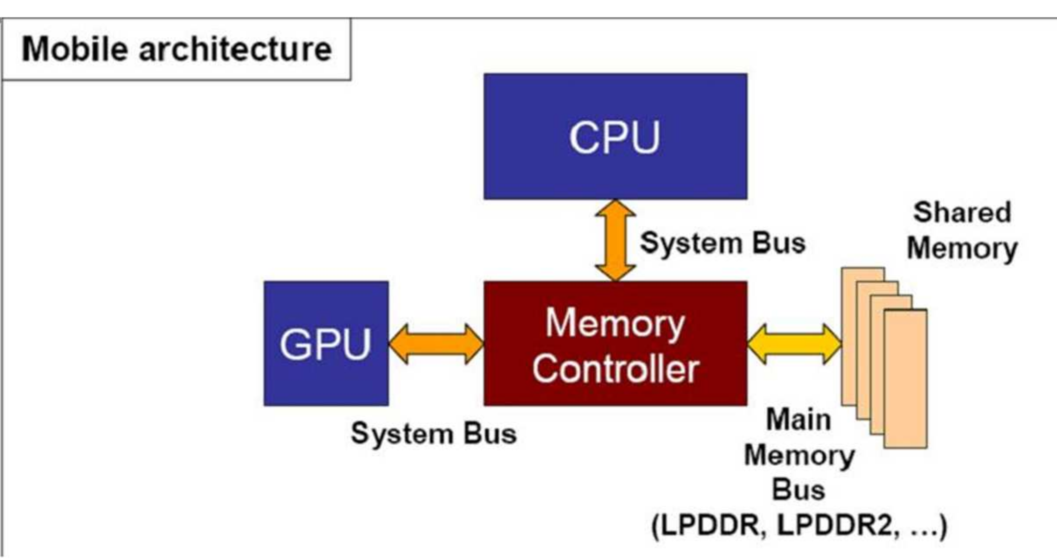
Outline

- Introduction
- 3D TSV technology choices
  - 3D-IC toolbox
  - Design perspectives
- Improving memory bandwidth
  - The memory link bottleneck
  - Overview of memory interfaces: bandwidth, power consumption, package integration
- Wide IO interface integration: Mag3D circuit
  - Wide IO technology
  - Mag3D circuit: architecture, Wide IO DRAM integration, design implementation challenges and current results
- Conclusion & perspectives

© CEA. All rights reserved

The memory link bottleneck



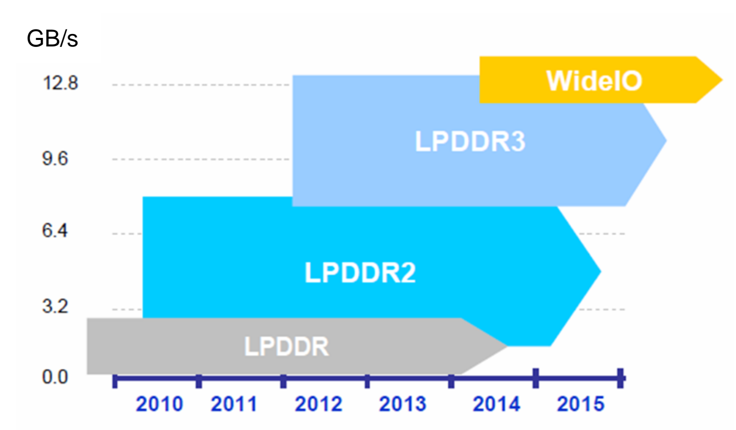


- The transfer Bandwidth between Memory and processor always been an issue for computing:
  - BW = nb of bits x bit rate
- The needs of computing:
  - 2014: 1TB/s
- For Mobile Computing architecture is more constraining, both Graphic and system memory on same support
- The needs of Mobile computing:
  - 2010: 4GB/s
  - 2012: 12GB/s
  - 201X: 100GB/s

© CEA. All rights reserved

Standardization evolves to 3D stacked memory

- JEDEC JC-42.6 Subcommittee for Low Power Memories:
  - LPDDR2, LPDDR3, WideIO
- LPDDR2:
  - Advanced power management features
  - Shared interface for nonvolatile memory (NVM) and volatile memory (SDRAM)
- LPDDR3:
  - Extension of LPDDR2
  - Bandwidth reaching 6.4GBps and allowing 12.8GBps for a dual channel configuration.
  - Supports POP packaging type
- WideIO:
  - Breakthrough technology
  - Uses chip-level three dimensional (3D) stacking with TSV interconnects
  - Bandwidth is 12.8GBps



Information from JEDEC web site

© CEA. All rights reserved

IO and power cost for a 100GB/s memory bandwidth

Memory link, peak bandwidth and power consumption efficiency		Cost for 100 GB/s memory bandwidth	
		Number of data IO pins	Interface power consumption
Multi-core SoC → LPDDR2 → DRAM 533 MHz I/O bus clock, 32 bits, 1.2 V, Double Data Rate	4.264 GB/s ~20 mW/Gb/s	770	16 W
Multi-core SoC → LPDDR3 → DRAM 800 MHz I/O bus clock, 32 bits, 1.2V, Double Data Rate	6.4 GB/s <20 mW/Gb/s	770	<16 W
Multo-core SoC → Wide I/O → DRAM 200 MHz I/O bus clock, 512 bits, 1.2 V, Single Data Rate	12.8 GB/s 4 mW/Gb/s	4100	3.2 W

- 3D integration of DRAM memory on processor with wide data interconnects provides higher data throughput and lower power consumption than 2D interfaces.

© CEA. All rights reserved

Memory-interface package evolution

DDR3

Off-Package

LPDDR2 / LPDDR3

POP Package

Wide I/O DRAM

3D IC with TSVs

Multi-core SoC

More integrated
Reduced power consumption
Higher bandwidth

© CEA. All rights reserved

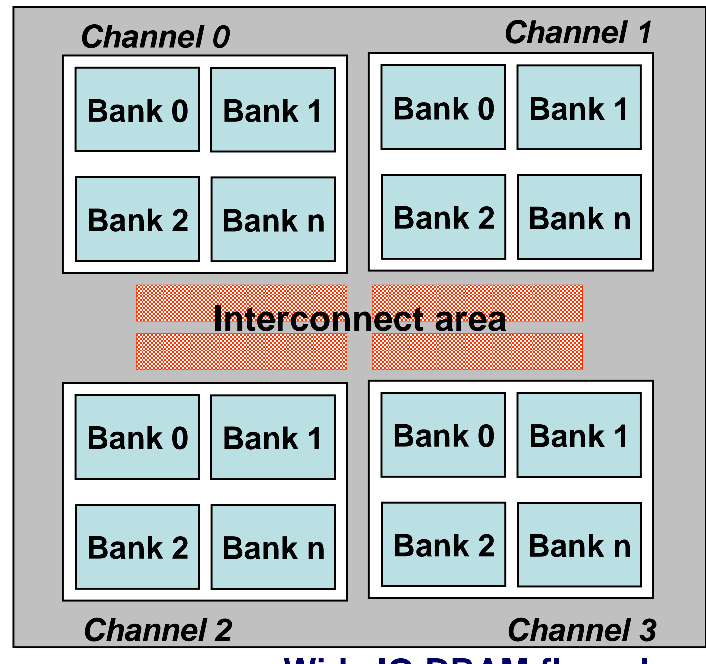
Outline

- Introduction
- 3D TSV technology choices
  - 3D-IC toolbox
  - Design perspectives
- Improving memory bandwidth
  - The memory link bottleneck
  - Overview of memory interfaces: bandwidth, power consumption, package integration
- Wide IO interface integration: Mag3D circuit
  - Wide IO technology
  - Mag3D circuit: architecture, Wide IO DRAM integration, design implementation challenges and current results
- Conclusion & perspectives

© CEA. All rights reserved

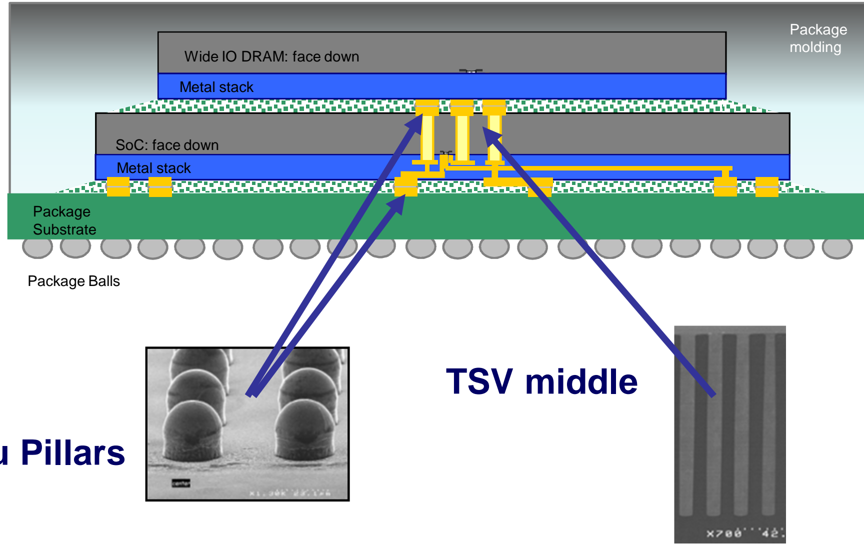
Wide IO technology

- Low power DRAM technology
- 12.8GByte/s peak bandwidth
- ~500mW power
- Quad-channel interface with large data bus operating at low frequency:
  - 128 bit wide data bus
  - Single Data Rate sampling mode @200MHz
  - Separate power-down & self-refresh per channel
- Timing performances:
  - Same as SDR 200MHz
- 1.2V low voltage supply
- Total I/O count including supply is in the range of ~1200:
  - IO physical location standardized at Jedec
  - 4 Arrays of 50x6; Pitch: 50µm x 40µm
  - Placed in the center area of the memory die
- Boundary scan test mode to detect bump connection failure between chips.
- Maximum number of stacked dies is 4



Wide IO DRAM floorplan

3D-IC with Wide IO DRAM



© CEA. All rights reserved

Mag3D test chip program

- Partnership between LETI, STEricsson, STMicroelectronics and Cadence
- Wide IO DRAM die is stacked on top of the SoC (Mag3D) in the same package
- Face to Back stacking
- Memory supplies and interconnect signals go through the SoC by means of TSVs

Technology assumptions	
Assembly	D2D
Stacking	F2B
TSV process	Via Middle
TSV density	10µm diameter
TSV xy pitch	50µm x 40 µm
Copper Pillars	25µm diameter

Package molding
Wide IO DRAM: face down
Metal stack
Central array of µbumps
1250 µ-bumps
SoC: face down
VDDQ supply
signal
supply
1250 TSVs
Metal stack
Peripheral bumps
Central matrix
985 bumps
Package Substrate
SoC signal and supply IOs, SDRAM muxed test IOs
µ-buffer supply IOs
SDRAM supply IOs
SoC signal and supply IOs, SDRAM muxed test IOs
Package Balls
475 balls

© CEA. All rights reserved

Wide IO integration into Mag3D

- 3D test chip baseline is the 65nm LETI MAGALI SoC
- The NoC architecture has been extended to support Wide IO memory
- Four independent data traffic and memory controllers have been added

Wide IO – Smart Memory Engine

TX_BIT tx_bit00n
MEPHISTO mep_01n
MEPHISTO mep_02n
TRX_OFDM trx_ofdm_03n
3D (ftol)
SME sme_01
3D (serial2)
SME sme_03
ARM11 arm11_00w
00 01 02 03
TRX_OFDM trx_ofdm_03e
TEST 3DNoC
SME_EXT sme_10
SME_WIDEIO sme_11
SME_WIDEIO sme_12
UDECASIP asip_13
10 11 12 13
nocif1
TEST Wide IO
UDECASIP asip_13
SME_WIDEIO sme_21
SME_WIDEIO sme_22
RX_BIT
nocif2
20 21 22 23
3D (serial2r)
SME sme_31
3D (normal)
SME sme_33
MEPHISTO_ HEATER mep_30w
30 31 32 33
MEPHISTO mep_33e
TRX_OFDM trx_ofdm_30s
TRX_OFDM trx_ofdm_31s
MEPHISTO mep_32s
MEPHISTO mep_33s

NIF | Data traffic & Channel 0 Controller | Test | µ-Buffer ESD | Channel 0
NIF | Data traffic & Channel 1 Controller | Test | µ-Buffer ESD | Channel 1
NIF | Data traffic & Channel 2 Controller | Test | µ-Buffer ESD | Channel 2
NIF | Data traffic & Channel 3 Controller | Test | µ-Buffer ESD | Channel 3

Wide IO SDRAM 4x128-bit channel - 166Mhz

© CEA. All rights reserved

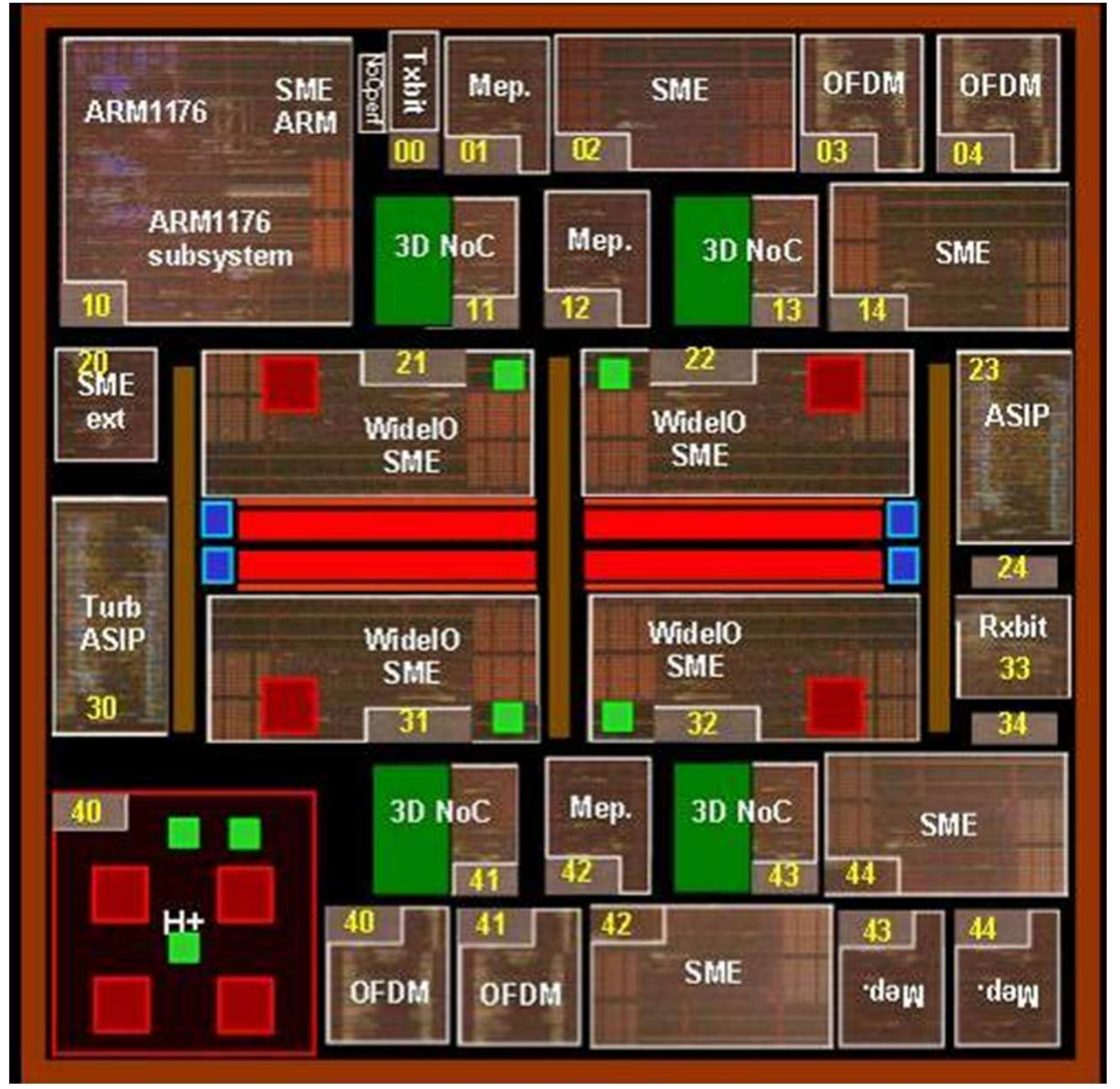
Wide IO SME: architecture

- Data traffic and memory controller
- WideIO PHY interface (STEricsson):
  - 166 MHz SDR (next 200MHz)
  - 128 bit data
  - ~1200 TSVs
  - 1.2V
- Memory controller (3rd party Wide IO controller)
  - 128 bit SDR
- Smart Memory Engine (LETI):
  - Data transfer handling between Wide IO, SRAM and ANoC
  - Integration within ANoC
  - Up to 3.2GB/s data bandwidth
- Specific design for Testability integration:
  - Direct access,
  - Boundary scan
  - MBIST

Wide IO DRAM Channel #n
128-bit data @ 166 MHz SDR
JTAG Port
SoC TAP
Test CTRL
DA, BS, Test Mux
128-bit data
PHY
capt_clk 166MHz
Quad div
PLL
Input clock
DFI: 128-bit data
Test clocks
Heater
Thermal sensor
Wide IO Memory Controller & MBIST (3rd Party IP)
clk
dram_clk 166MHz
Reg. port
AXI regACLK
AXI Interface
AXI3
axi0_CLK
Static mux
ST ERICSSON
sys_clk, 400MHz
128-bit data
128-bit data @ 400 MHz
Read-write decoder
AXI port
Config
configuration & data mover
RAM 0
RAM 1
RAM 2
RAM 3
Output port
Output Com. Control.
Data reader
µpg cache
Memory arbiter
Network On Chip
Input port
Input Com. Control.
Data writer
Reset
sys_clk 400 MHz
Local Clock Gen
Reference clock
Base_addr
size
wr/rd_ptr
...
leti

© CEA. All rights reserved

Mag3D topology & floorplan



Legend

PLL

Lite PHY location

WideIO µ-bumps

Temp. sensor

Heater

3D NoC µ-bumps

aNoC routing channels

MAG3D few numbers :

-CMOS 65nm + TSV middle Ø10µm

-8500µm x 8500µm (72mm2)

-1.8 M-instances

-400 Macros

-270 IO pads

-985 Bumps (Flip-Chip)

-1980 TSV for 3D NoC

-1250 TSV for WideIO memory

Performances :

-Units in the [350 - 400] MHz range

-Asynchronous NoC ~ 550 MHz

-Overall power : 500mW – 1W

© CEA. All rights reserved

Physical implementation challenges with 3D design

- Supply distribution for Wide IO memory and micro-buffers
- 3D interconnect routing: flip chip bumps, RDL, TSV and backside micro-bumps
- ESD strategy
- Test constraint integration
- Final verification (LVS for 3D stack)

© CEA. All rights reserved

Conclusion & Perspectives

- TSV technology is mature enough to allow coarse grain partitioning with ~500 TSV per mm^2
- In mobile computing, off-package memory interfaces have reached their limit above ~10GByte/s:
- In the coming years, multi-core processor will require more than 100GB/s for a reasonable power budget
- 3D stacking technology enables a power efficiency breakthrough in memory interconnect
- Progress to be done with 3D design flow

→ Be ready to integrate 3D in your current architecture hypothesis

© CEA. All rights reserved

Many Thanks …

- To our partners in this project
  - STMicroelectronics, ST-Ericsson, CADENCE

© CEA. All rights reserved

Save the date now!

→ June 19-20, 2012

MINATEC - GRENOBLE - FRANCE

Thank you for your attention

Leti

LABORATOIRE D'ÉLECTRONIQUE
ET DE TECHNOLOGIES
DE L'INFORMATION

CEA-Leti
MINATEC Campus, 17 rue des Martyrs
38054 GRENOBLE Cedex 9
Tel. +33 4 38 78 36 25

www.leti.fr