A 0.9 pJ/bit, 12.8 GByte/s WideIO Memory Interface in a 3D-IC NoC-based MPSoC.

Denis Dutoit[1], Christian Bernard[1], Séverine Chéramy[1], Fabien Clermidy[1], Yvain Thonnart[1], Pascal Vivet[1], Christian Freund[2], Vincent Guérin[2], Stéphane Guilhot[2], Stéphane Lecomte[2], Gianni Qualizza[2], Julien Pruvost[3], Yves Dodo[3], Nicolas Hotelier[3], Jean Michailos[3]

[1] CEA, Leti, Minatec Campus, Grenoble, France; [2] ST-Ericsson, Grenoble, France; [3] STMicrolectronics, Grenoble, Crolles, France

E-mail: *denis.dutoit@cea.fr*

Abstract

3D Integrated Circuit (3D-IC) opens architecture opportunities for improved SoC-to-memory interconnect bandwidth between dies. This paper presents the design of a two-tier 3D-IC composed of one NoC-based MPSoC and one multi-channel WideIO mobile SDRAM stacked in a face-to-back configuration. Measurements of the 3D-IC show that the targeted 12.8 GByte/s bandwidth is achieved in worst case conditions, while offering a 0.9 pJ/bit 3D I/O link power efficiency.

Introduction

Real-time and high-performance processing is typically implemented as semi-heterogeneous Multi-Processor System-on-Chip (MPSoC) [1][2] where multiple type of Intellectual Property cores share high capacity, low power and high speed off-chip memory. The new generation of mobile applications as High Definition TV, gaming, augmented reality or Software Defined Radio will lead to complex memory communication use cases with heavy demand in data bandwidth while keeping limited power consumption. Three-dimensional (3D) stacking technology, which enables low-latency, high bandwidth and very dense die-to-die interconnects, is the right opportunity to develop new vertical connection schemes for memory chip stacked upon a MPSoC [3][5]. This work proposes and realizes a novel approach for MPSoC to 3D-stacked-memory communication by introducing a multi-channel WideIO interface and traffic manager in a semi-heterogeneous Network-on-Chip (NoC) based architecture [1].

Architecture

The MPSoC architecture (see Fig. 5 circuit floorplan) is organized around a 16-router Asynchronous NoC backbone on which 24 programmable units are connected. High-performance computing comes from both the processing cores (MEP) and the dedicated operators (OFDM) while the Data and Configuration Memory (DCM) unit [4] manages intra-chip NoC data flows between operators and on-chip memories.

A 3D capable version of the DCM called DCM-WIDEIO (see Fig. 1) has been designed to act as an advanced memory traffic manager for each channel of the new standardized WideIO mobile DRAM (JEDEC: http://www.jedec.org). The DCM-WIDEIO unit is replicated 4 times as fully independent hard macros to fit with the logical and physical 3D interface footprint of the quad channel WideIO memory. In addition to its legacy task of data manipulation among the MPSoC processing operators, each DCM-WIDEIO unit offers 3.2 GByte/s full-rate data flow per 128-bit, 200MHz DRAM channel and low-latency medium-bandwidth data manipulation between NoC and WideIO.

The 128-bit data memory requests sent by the Traffic Controller are translated into DRAM logical protocol with the Memory Controller and then into DRAM timing compliant signals through the physical layer interface (PHY details depicted in Fig 1).

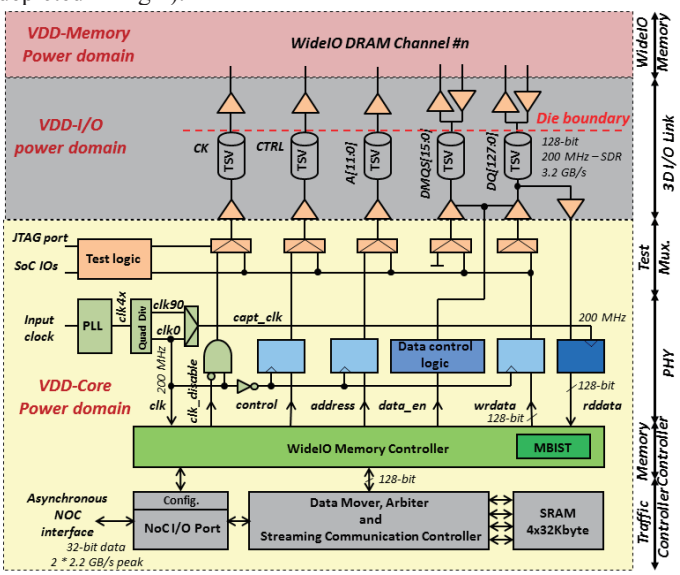


Fig. 1. DMC-WIDEIO with PHY details and power domains.

The PHY consists in a 128-bit wide DRAM data capture interface and operates up to 200 MHz in Single Data Rate (SDR) mode, offering an aggregate 12.8 GB/s memory link. With SDR protocol at 200MHz frequency, the WideIO PHY solution is less constrained than traditional LPDDR2 or LPDDR3 architectures running up to 800MHz. It allows simple digital std-cell based design in favor of ultra-low power consumption by removing all power hungry programmable Delay-Locked Loop devices (DLL). In our architecture, a DLL-less fixed quarter cycle delayed clock has been implemented to avoid set-up time violation during data capture at high frequency.

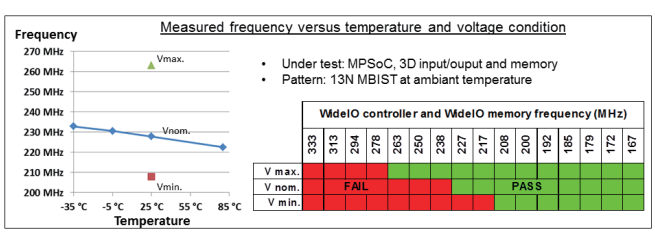


Fig. 2. Frequency characterization.

In Fig. 2, the Si measured 208 MHz operating frequency at all minimum voltage values (Vmin) with a high activity 13N MBIST test pattern proves the capability and robustness of our

technology and design solutions as predicted by our simulations. Moreover, applying maximum voltage values, the performance of a 266 MHz WideIO at 17 GB/s has been achieved.

To ensure the integrity and robustness of 3D signals through the heterogeneous die stack, micro-buffers have been inserted and placed between the TSVs as close as possible from each input/output (I/O) signal. The micro-buffer cell embeds electrostatic discharge (ESD) protection required for die assembly process. As shown on Fig. 1 and as described in Fig. 4, classical Design-for-Test (DFT) technique such as TAP based scan has been applied to 3D signals for connection failure detection, but with some 3D specific add-ons like the SoC embedded Memory Built-in-Self-Test (MBIST) with bitmap capability for memory failure analysis after stack assembly. All DFT features are accessible from the SoC JTAG port and managed by an IEEE1500 TAP controller.

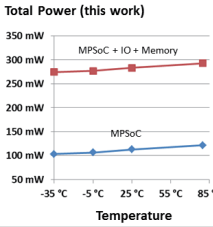
Technology and Performances

Fig. 4 and 5 show respectively the main process features and the 3D 2-layer stack Integrated Circuit (3D-IC). The memory is stacked on top of the MPSoC in a Face-to-Back configuration. The interconnect, located in the center area of the memory die, consists of an array of micro-bumps with a 40 µm x 50 µm pitch, compatible with the JEDEC standard. With this configuration, TSVs are required within the MPSoC for WideIO memory supply distribution and signal propagation. The MPSoC is packaged with a flip-chip technology to improve the power integrity of the 3D stack. The 3D-IC is arranged within a 12 x 12 ball grid array package with a 0.4 mm ball pitch. The substrate comprises four layers High Density Interconnect (HDI), with copper filled vias stacked between each layer.

Power Consumption

The MBIST test infrastructure embedded in the SoC and exercised at full speed in parallel for the four channels has been used in an Automated Test Equipment (ATE) environment for accurate 3D stack power measurement. During the peak activity section of this test pattern, we have monitored the power consumption of the complete 3D stack at different temperatures (from -35°C to +80°C as shown in Fig. 3). The total power measured on the 3D-IC is 293mW at 80°C, including three contributions: the MBIST peak activity on the MPSoC traffic manager, the TSV based 3D I/O link (12.8 GB/s) and the WideIO memory plane. It represents a power efficiency of 2.86 pJ/bit for the full WideIO system. When we remove the SoC and memory part of the power consumption to only consider the TSV based 3D I/O interconnect, the power consumption becomes 91mW. The 3D I/O link power efficiency is then 0.9 pJ/bit which represents an improvement of a factor 4 with state-of-the-art LPDDR3 off-package I/O link (3.7 pJ/bit with 5pF load in [6]).

Measured power consumption versus temperature and comparison with LPDDR3



Memory Type	LPDDR3 [6]	WideIO This work
Package	PoP / Discrete	3D-IC
BW (Gbyte/s)	6.4 GB/s	12.8 GB/s
Total power (MPSoC + I/O + memory)		293 mW(1)
MPSoC power (VDD-Core in Fig. 1)		121 mW(1)
Memory Power (VDD-Mem. in Fig. 1)		81 mW(1)
I/O power (VDD-I/O in Fig. 1)		91 mW(1)
I/O power efficiency	3.7 pJ/bit(2)	0.9 pJ/bit(1)

Measurements conditions: (1) 13N MBIST, 80°C
(2) Read, 5pF load without ODT

Fig. 3. Power consumption and comparison with LPDDR3.

Conclusion

3D stacking is a unique opportunity enabling memory-interconnect evolution to higher bandwidth with a limited power impact thanks to very short connections between dies. As evidence, this chip is demonstrating a WideIO memory assembly with flip-chipped MPSoC allowing a gain of a factor of 4 in mobile memory interface energy efficiency.

References

[1] F. Clermidy, et al. "A 477mW NoC-Based Digital Baseband for MIMO 4G SDR", IEEE International Solid-State Circuits Conference (ISSCC), Feb. 2010, pp. 278-279.

[2] D. Melpignano, L. Benini, E. Flamand, et al. "Platform 2012, a many-core computing accelerator for embedded SoCs: Performance evaluation of visual analytics applications", Design Automation Conference (DAC), June 2012.

[3] Dae Hyun Kim, et al. "3D-MAPS: 3D Massively Parallel Processor with Stacked Memory", IEEE International Solid-State Circuits Conference (ISSCC), Feb. 2012, pp. 188 – 189.

[4] J. Martin, et al. "A Microprogrammable Memory Controller for high-performance dataflow applications", European Solid Sate Circuit Conference (ESSCIRC), Sept. 2009.

[5] Jung-Sik Kim, et al. "A 1.2V 12.8GB/s 2Gb mobile Wide-I/O DRAM with 4×128 I/Os using TSV-based stacking", IEEE International Solid-State Circuits Conference (ISSCC), Feb. 2011, pp. 496 – 498.

[6] Yong-Cheol Bae, et al. "A 1.2V 30nm 1.6Gb/s/pin 4Gb LPDDR3 SDRAM with input skew calibration and enhanced control scheme," Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International, pp.44-46, 19-23 Feb. 2012.

MPSoC Architecture	NoC Architecture	32-bit, asynchronous, 550 Mflit/s, fifo-based GALS interfaces, Local Clock Generators
	Units	9 memory units including 4 DCM WideIO controllers, 14 processing RX+TX units
	Host processor	ARM core & peripherals (Embedded Trace Module, Interrupt cont., Timer, Uart)
Design-for-Test	Test Modes: *TAP* *Boundary Scan* *Memory BIST* *Direct Access* *Stuck-at, transition* *PLL test*	Test Description: To test TAP controller To test TSV+Copper-Pillar connections between SoC die & memory die To test the whole memory die from the SoC die with a 13N memory BIST pattern To generate direct memory accesses for debug purpose To test the memory controller logic, using standard scanable FFs To test the PLL (generating the WideIO DRAM clock, for each memory channel)
	Test Architecture	IEEE 1500 SOC TAP controller + one memory BIST per WideIO memory channel
	Test Features	Coverage: fault 86%- Memory BIST 100%- PLL 100%, MBIST at speed
Technology	CMOS Process	65nm STMicroelectronics, Low-Power, Multi-VT (LVT, SVT, HVT)
	3D Process	1016 TSV middle (diam 10 µm, aspect ratio 8:1) & Copper Pillar (diam 20 µm), 50µm x 40µm pitch, no backside RDL, Face2Back stacking, Die2Die assembly
	Substrate	flip-chip assembly, 933 frontside bumps, with 150 µm pitch
	Package	12x12x1.2 BGA, 459 balls, with 0.4 mm ball pitch
	Design complexity	15 Mgates / 1.63 Mbytes SRAM / Total 228 Mtransistors / 276 IO pads
	Die Size / Core Size	8.5mm * 8.5mm = 72.2 mm² / 8.2mm * 8.2mm = 67.2 mm²
	Supply Voltage	SoC: 1.2 V Core, 1.2V 3D I/O, 2.5 V 2D I/O pads, Memory: 1.2 V WideIO memory core1, 1.8 V WideIO memory core2, 1.2V 3D I/O.

Fig. 4. 3D Integrated-Circuit main features.

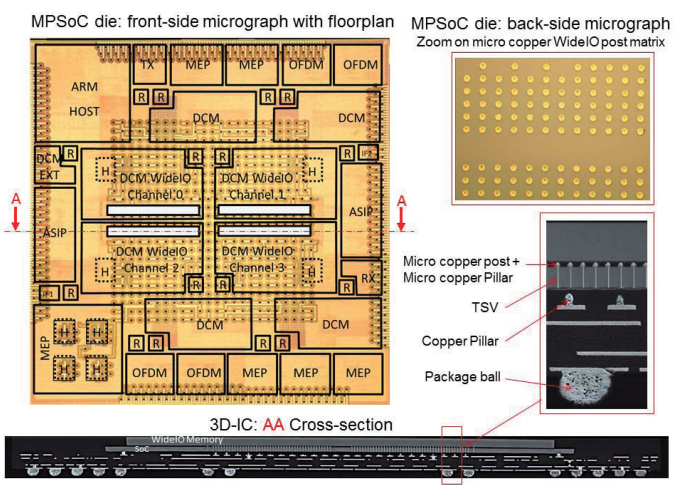


Fig. 5. 3D Integrated-Circuit microphotographs.