

INTACT: A 96-Core Processor With Six Chiplets 3D-Stacked on an Active Interposer With Distributed Interconnects and Integrated Power Management

Pascal Vivet^{ID}, Member, IEEE, Eric Guthmuller, Yvain Thonnart, Member, IEEE,
 Gael Pillonnet^{ID}, Senior Member, IEEE, César Fuguet, Ivan Miro-Panades^{ID}, Member, IEEE,
 Guillaume Moritz, Jean Durupt, Christian Bernard, Didier Varreau, Julian Pontes, Member, IEEE,
 Sébastien Thuries, David Coriat, Michel Harrand, Member, IEEE, Denis Dutoit, Didier Lattard,
 Lucile Arnaud, Jean Charbonnier, Member, IEEE, Perceval Coudrain, Arnaud Garnier,
 Frédéric Berger, Alain Gueugnot, Alain Greiner, Quentin L. Meunier,
 Alexis Farcy, Alexandre Arriordaz, Séverine Chéramy^{ID}, Member, IEEE,
 and Fabien Clermidy, Member, IEEE

Abstract—In the context of high-performance computing, the integration of more computing capabilities with generic cores or dedicated accelerators for artificial intelligence (AI) application is raising more and more challenges. Due to the increasing costs of advanced nodes and the difficulties of shrinking analog and circuit input output signals (IOs), alternative architecture solutions to single die are becoming mainstream. Chiplet-based systems using 3D technologies enable modular and scalable architecture and technology partitioning. Nevertheless, there are still limitations due to chiplet integration on passive interposers—silicon or organic. In this article we present the first CMOS active interposer, integrating: 1) power management without any external components; 2) distributed interconnects enabling any chiplet-to-chiplet communication; and 3) system infrastructure, design-for-test, and circuit IOs. The INTACT circuit prototype integrates six chiplets in FDSOI 28-nm technology, which are 3D-stacked onto this active interposer.

Manuscript received June 11, 2020; revised September 17, 2020 and October 27, 2020; accepted November 2, 2020. Date of current version December 24, 2020. This paper was approved by Guest Editor Dejan Markovic. This work was supported in part by the French National Program Programme d'Investissements d'Avenir, IRT Nanoelec under Grant ANR-10-AIRT-05, in part by the SHARP CA109 CATRENE Project, in part by the MASTER3D CT312 CATRENE Project, and in part by the Hubeo+ CARROT Project. (*Corresponding author: Pascal Vivet*)

Pascal Vivet, Eric Guthmuller, Yvain Thonnart, Gael Pillonnet, César Fuguet, Ivan Miro-Panades, Guillaume Moritz, Jean Durupt, Sébastien Thuries, David Coriat, Michel Harrand, Denis Dutoit, Didier Lattard, Lucile Arnaud, Jean Charbonnier, Perceval Coudrain, Arnaud Garnier, Frédéric Berger, Alain Gueugnot, Séverine Chéramy, and Fabien Clermidy are with CEA, University Grenoble Alpes, 38054 Grenoble, France (e-mail: pascal.vivet@cea.fr).

Christian Bernard and Didier Varreau, retired, were with CEA Grenoble, 38054 Grenoble, France.

Julian Pontes was with CEA Grenoble, 38054 Grenoble, France. He is now with ARM, Sheffield S1 4LW, U.K.

Alain Greiner and Quentin L. Meunier are with LIP6 Lab, University Paris Sorbonne, 75252 Paris, France.

Alexis Farcy is with STMicroelectronics, 38920 Crolles, France.

Alexandre Arriordaz is with Mentor, A Siemens Business, 38330 Montbonnot, France.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JSSC.2020.3036341>.

Digital Object Identifier 10.1109/JSSC.2020.3036341

in 65-nm process, offering a total of 96 computing cores. Full scalability of the computing system is achieved using an innovative scalable cache-coherent memory hierarchy, enabled by distributed network-on-chips, with 3-Tbit/s/mm² high bandwidth 3D-plug interfaces using 20-μm pitch micro-bumps, 0.6-ns/mm low latency asynchronous interconnects, while the six chiplets are locally power-supplied with 156-mW/mm² at 82%-peak-efficiency dc-dc converters through the active interposer. Thermal dissipation is studied showing the feasibility of such approach.

Index Terms—3D technology, active interposer, chiplet, network-on-chip (NoC), power management, thermal dissipation.

I. INTRODUCTION

IN THE context of high-performance computing (HPC) and big-data applications, the quest for performance requires modular, scalable, energy-efficient, low-cost many-core systems. To address the demanding needs for computing power, system architects are continuously integrating more cores, more accelerators and more memory in a given power envelope [1]. It appears that similar needs and constraints are emerging for the embedded HPC domain, in transport applications for instance with autonomous driving, avionics, and so on.

All these application domains require highly optimized and energy-efficient functions: generic ones such as cores, GPUs, embedded FPGAs, dense and fast memories, and also more specialized ones, such as machine learning and neuro-accelerators to efficiently implement the greedy computing demand of Big Data and artificial intelligence (AI) applications.

Circuit and system designers are in need of a more affordable, scalable, and efficient way of integrating those heterogeneous functions, to allow more reuse, at circuit level, while focusing on the right innovations in a sustainable manner. Due to the slowdown of advanced CMOS technologies (7 nm and below), with yield issues, design, and mask costs,

the innovation and differentiation through single die solution is not viable anymore. Mixing heterogeneous technologies using 3D is a clear alternative [2], [3]. Partitioning the system into multiple chiplets 3D-stacked onto large-scale interposers—organic substrate [4], 2.5D passive interposer [5], or silicon bridge [6]—leads to large modular architectures and cost reduction in advanced technologies using a Known Good Die (KGD) strategy and yield management.

Nevertheless, the current passive interposer solutions still lack flexible and efficient long-distance communication, smooth integration of chiplets with incompatible interfaces, and easy integration of less-scalable analog functions, such as power management and system input output signals (IOs). We present the first CMOS Active Interposer, measured on silicon, integrating power management, distributed interconnects, enabling an innovative scalable cache-coherent memory hierarchy. Six chiplets are 3D-stacked onto the active interposer, offering a total of 96 cores.

The outline of this article is as follows. Section I introduces the chiplet paradigm in more detail, with a state of the art on these technologies and the proposed concept of active interposer. Section II presents an overview of the INTACT demonstrator architecture and 3D technology, while Sections III–VIII detail the various sub-elements of the circuit: computing chiplet, power management, distributed interconnects, and testability. Section IX addresses the thermal issues. Finally, Sections X and XI present the final circuit results and conclusion.

II. CHIPLET AND ACTIVE INTERPOSER PRINCIPLE

A. Chiplet Partitioning: Concept and Challenges

Chiplet partitioning is raising new interest in the research community [7], in large research programs [8], and in the industry [9]. It is actually an idea with a long history in the 3D technology field [2]. The concept of chiplet is rather simple: divide circuits in modular sub-systems, in order to build a system as an LEGO-based approach, using advanced 3D technologies.

The motivation for chiplet-based partitioning is as follows.

- It is driven by cost. Due to increasing issues in advanced CMOS technologies (7 nm and below), achieving high yield on large dies in acceptable costs is not possible anymore, while shrinking all the analog intellectual property blocks (IPs) (power management, Fast IO SerDes, and so on) is becoming increasingly difficult. By dividing a system into various sub-modules, called chiplets, it is possible to yield larger systems at an acceptable cost, thanks to KGD sorting [10].
- It is driven by modularity. By an elegant divide and conquer partitioning scheme, *chipletization* allows to build modular systems from various building blocks and circuits, focusing more on functional aspects than on technology constraints. Circuit designers can deeply optimize each function: generic CPUs, optimized GPUs, embedded FPGAs, dedicated accelerators for machine learning, dense memory solutions, IO, and services, while the system designer is picking the best combination to build a differentiated and optimized system.

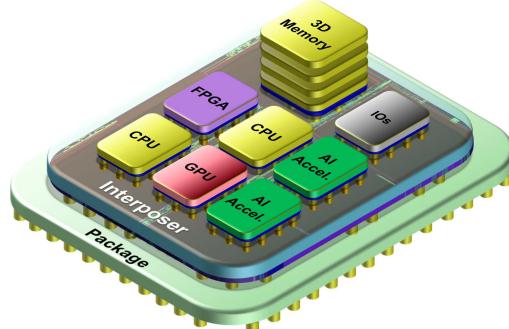


Fig. 1. Chiplet partitioning concept.

- It is enabled by heterogeneous integration. For chiplets, the right technology is selected to implement the right function: advanced CMOS for computing, DRAM for memory cubes like high bandwidth memory (HBM) [11], non-volatile memory (NVM) technology for data processing within AI accelerators [12], mature technology for analog functions (IOs, clocking, power management, and so on). Chiplet integration is then performed using advanced 3D technologies, which are getting more and more mature, with reduced pitches, using through-silicon via (TSV) and micro-bumps [5] or even more advanced die-to-wafer hybrid bonding technologies [3], [13].

To benefit from all these advantages and possibilities, there are nevertheless clear challenges for chiplets. The ecosystem needs to change progressively from IP-reuse to chiplet-reuse; this requires fundamental changes in the responsibilities of the various providers. These constraints are economical rather than technical, but they are strongly driving the technical choices.

For system-level design, the simple LEGO cartoon (Fig. 1) needs some adequate methodologies to address system modeling, cost modeling, to perform technology and architecture partitioning while achieving an optimized system. A strong movement is building momentum toward the standardization of chiplet interfaces to enable this modularity between various vendors [14].

Finally, many circuit level design issues arise: design of energy-efficient chiplet interfaces, testability, power management, and power distribution, final system sign-off in terms of timing, power, layout, and reliability, thermal dissipation. To address these 3D design challenges, new CAD tools and associated design flows must be developed [49].

In this article, a partitioning using identical chiplets is proposed to scale-out a large distributed computing system offering 96-cores, by using heterogeneous technologies. Many circuit design aspects are addressed in terms of chiplet interfaces, distributed interconnects, power management, testability, and associated CAD flows.

B. State of the Art on Interposers

In order to assemble the chiplets together, various technologies have been developed and are currently available in the industry (Fig. 2).

Firstly, organic substrate is the lowest cost solution, while offering larger interconnect pitches ($130 \mu m$). This technology has been adapted by AMD for their EPIC

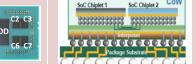
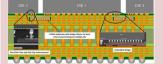
| Organic Substrates | Passive interposer (2.5D) | Silicon bridges | 3D Vertical Stacking |
|---|---|---|---|
|  |  |  |  |
| AMD, 4-chiplet ISCC 2018 | AMD, 4-chiplet ISCC 2020 | TSMC, CoWoS, VLSI 2019 | INTEL, EMIB bridge, ISCC 2017 |

Fig. 2. State-of-the-art on recent interposer and 3D technologies.

processor family, with the first version with up to four chiplets [4] and a recent version with up to eight chiplets using a central IO die to distribute the system level interconnects [15]. Passive interposers, also called 2.5D integration, as proposed for instance by TSMC CoWoS [5] enable more aggressive chip-to-chip interconnects and pitches ($40 \mu\text{m}$) but are still limited to “wire only” connections. A trade-off in terms of cost, pitches, and performances can be achieved by using a silicon bridge embedded within the organic substrate as presented by INTEL and their EMIB bridge [6]. Finally, regular 3D stacking (for vertical assembly) may also be used, which is also orthogonal and complementary of interposer approaches. INTEL has presented recent results with their Foveros technology and Lakefield processor [16].

All these solutions are promising and show clear benefits in terms of cost and performances. Nevertheless, various challenges still arise.

- Inter-chiplet communication is mostly limited to side-by-side communication, due to wire-only interposers. Longer range communication should rebound in the chiplets themselves, which is not scalable to build larger systems with numerous chiplets. The recent solution from AMD with their input output die (IOD) [15] is partially solving these issues, with better communication distribution and easier IO integration, but may still not scale further on the long term.
- Current interposer solutions do not integrate themselves with less scalable functions, such as IOs, analogs, power management, close to the chiplets. The recent solution from INTEL with digital on top of analog partitioning is solving this issue, but is still limited today to a single die [16].
- Finally, it is currently complex to integrate chiplets from different sources, due to missing standards, even if strong initiative is on-going [8], [14]. Wire-only interposers prevent the integration of chiplets using incompatible protocols, while active interposer enables to bridge them, as adapted by zGLUE Inc. [51].

In order to tackle all these issues, this article presents the concept of Active Interposer, which integrates logic functions within the large interposer area. The concept has been already introduced before, either as a low cost and limited active-light solution for ESD integration [17] or with system level architecture explorations showing the capability to scale larger systems [19], [20]. Section II-C presents the active interposer concept, enabled by technology improvements [18].

C. Active Interposer Principle and Partitioning

The proposed active interposer concept is detailed in Fig. 3. Chiplets can be either identical or different, for homogeneous

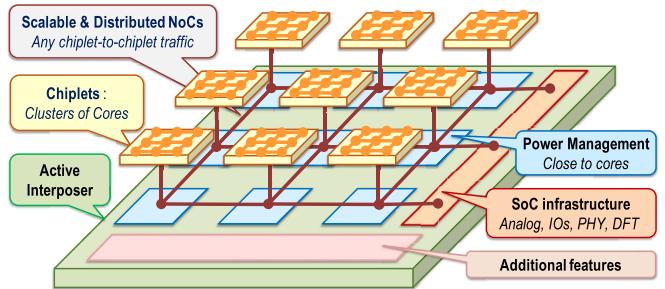


Fig. 3. Active interposer concept and main features.

functions as presented here, or differentiated functions, as presented in Fig. 1. Chiplets, implemented in an advanced CMOS technology, may themselves be composed of clusters of cores. Each chiplet can contain its own interconnects for intra-chip communication, which are extended in 3D for chiplet-to-chiplet communication. The CMOS interposer integrates a scalable and distributed network-on-chip (NoC), which offers the main capability of allowing any chiplet-to-chiplet traffic, without interfering with unrelated chiplets. As a conclusion, a hierarchical 3D NoC is obtained, with 2D NoC within the chiplet, 2D NoC within the active interposer, which can be further refined and differentiated according to traffic classes. Moreover, dense 3D interconnects enable high bandwidth density with parallel signaling. Such a communication scheme enables fully modular and scalable cache-coherent architecture, for offering large many-cores [20], [28], [29].

In order to provide efficient power supply to each chiplet, power management and associated power converters can be directly implemented within the active interposer, to bring power supply closer to the cores, for increased energy efficiency in the overall power distribution hierarchy, and allowing dynamic voltage and frequency scaling (DVFS) scheme at the chiplet level. Moreover, all the less-scalable functions, such as analog IPs, clock generators, and circuit IOs with SerDes and PHYs for off-chip communication, as well as the regular system-on-chip infrastructure, such as low performance IOs, test, debug, and so on, can also be implemented in the bottom die. Finally, additional features can be integrated into the active interposer, to specialize for a given application, enabling to differentiate the overall system. For instance, if incompatible chiplets are assembled, the active interposer can implement protocol bridges.

Due to the additional power budget within the interposer, the thermal challenge of 3D might increase. Nevertheless, most of the power budget is within the chiplets, thermal dissipation issues are then limited, as presented in Section IX.

Regarding technology partitioning, the active interposer should be implemented using a mature technology, with a low logic density to achieve high yield. Large logic density within a large interposer would lead, even using a mature technology, to an un-yieldable and costly system. A difference of at least two technology nodes between the computing chiplets and the interposer should lead to an acceptable cost, while allowing enough performances in the bottom die for analog and PHYs to sustain the overall system performances, as done in [16].

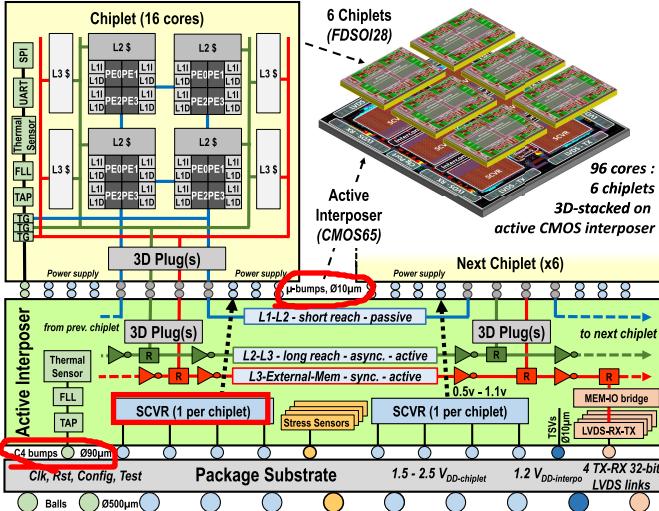


Fig. 4. INTACT overall circuit architecture.

III. INTACT CIRCUIT ARCHITECTURE

The proposed Active Interposer concept is implemented within a large-scale circuit demonstrator, offering 96-cores, called INTACT for “*Active Interposer*.”

A. IntAct Circuit Architecture

INTACT is the first CMOS active interposer [21] integrating: 1) switched capacitor voltage regulator (SCVR) for on-chip power management; 2) flexible system interconnect topologies between all chiplets for scalable cache coherency support; 3) energy-efficient 3D-Plugs for dense inter-layer communication; and 4) a memory-IO controller and PHY for socket communication. Fig. 4 presents an overview of the overall many-core circuit architecture, with the chiplets, the distributed interconnects, the integrated power management, and the system infrastructure, which are detailed hereinafter.

Each chiplet is a 16-core sub-system composed of four computing clusters of four cores, integrating their own distributed coherent memory caches, and their associated system level interconnects. The chiplet architecture and associated memory hierarchy is presented in Section IV.

The chiplet interconnects are extended through the active interposer for chiplet-to-chiplet communication using distributed NoCs and various kinds of communication links, using so-called “3D Plug” communication interfaces. For off-chip communication, the active interposer integrates a memory-IO controller and 4×32 bits 600-Mb/s bidirectional LVDS links offering a total of 19.2-GB/s off-chip bandwidth. The communication IPs and overall communication architecture are presented in Sections VI and VII, respectively.

The active interposer integrates a power management IP for supplying individually each chiplet and offering on-demand energy-efficient power management, below each chiplet and surrounded by pipeline NoC links. The SCVR is presented in more detail in Section V.

Finally, the active interposer integrates some regular System-on-Chip infrastructure elements such as clock and reset generation, thermal sensors, stress sensors, low speed interfaces (UART, SPI) for debug and configuration, and

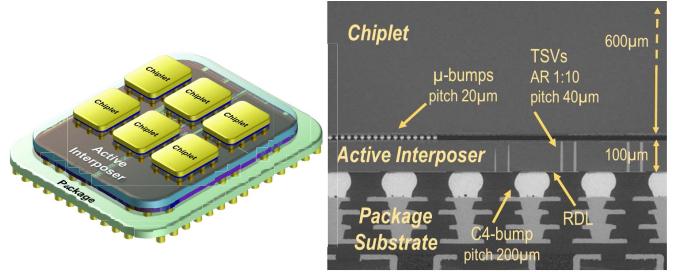


Fig. 5. INTACT: from concept to 3D-cross section.

design-for-test (DFT) logic for KGD sorting and final test. 3D testability challenges and associated DFT are presented in more detail in Section VIII.

In conclusion, INTACT offers a large-scale cache-coherent many-core architecture, with a total of 96 cores in six chiplets (four cores \times four clusters \times six chiplets), which are 3D-stacked onto the active interposer.

B. INTACT Circuit Technology Partitioning

The 22-mm² chiplets are implemented in a 28-nm FDSOI CMOS node, while the 200-mm² active interposer is using a 65-nm CMOS node, which is a more mature technology. As presented in Section III, this technology partitioning exhibits two technology node differences between the computing die and the active interposer. This enables enough performance in the bottom die for the interconnects, the analog parts, and the system IOs, while still allowing a yieldable large-scale active interposer.

Even though complex functions are integrated, the yield of the active interposer is high thanks to this mature 65-nm node and a reduced complexity (0.08 transistor/ μm^2 , see Section II-C), with 30% interposer area devoted to SCVR variability-tolerant capacitors scheme. This technology partitioning leads to a practical and reachable circuit and system in terms of silicon cost using advanced 3D technologies (more details in terms of yield analysis can be found in [22]).

C. INTACT Physical Design and 3D Technology Parameters

For enabling system integration, and allowing efficient chiplet-to-chiplet communication, an aggressive 3D technology has been developed and used. A summary of the respective chiplet, interposer, and 3D technologies is given in Table I.

As presented in Fig. 5 with the circuit 3D-cross section, the six chiplets are 3D-stacked in a face-to-face configuration using 20- μm -pitch micro-bumps (μ -bumps) onto the active interposer (2 \times smaller pitch compared to state of the art [23]). These dense chip-to-chip interconnects enable a high bandwidth density, up to 3 Tbit/s/mm² as detailed in Section VI-A, using parallel signaling through thousands of 3D signal interfaces. For bringing power supplies and allowing off-chip communication, the active interposer integrates TSV-middle with a pitch of 40 μm and an aspect ratio of 1:10 (10 μm diameter for a silicon height of 100 μm) and a keep-out zone of 10 μm . Finally, the overall system is assembled

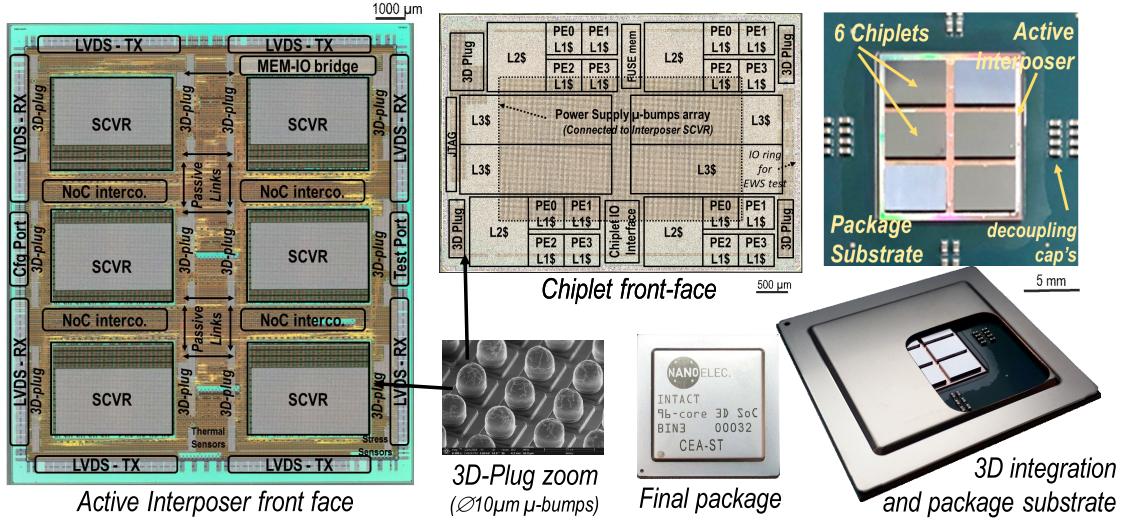


Fig. 6. Chiplet and active interposer floorplans, details of the 3D-plug μ -bumps, final 3D integration and package.

TABLE I
INTACT MAIN CIRCUIT FEATURES AND 3D TECHNOLOGY DETAILS

| Chiplet technology | FDSOI 28nm, 10 metals, 0.5V-1.3V+adaptive biasing |
|------------------------|---|
| Chiplet area | 4.0 mm x 5.6 mm = 22.4 mm ² |
| Chiplet complexity | 395 Million transistors, 18 transistors/ μm^2 density |
| Interposer tech. | CMOS 65nm bulk, 7 metals, MIM option, 1.2V |
| Interposer area | 13.05 mm x 15.16 mm = 197.8 mm ² |
| Interposer complexity | 15 Million transistors, 0.08 transistors/ μm^2 density |
| 3D technology | Face2Face, Die2Die assembly onto active interposer |
| μ -bump technology | $\varnothing 10\mu\text{m}$, pitch 20 μm |
| # μ -bumps | 150 000 (20k signals + 120k powers + 10k dummies) |
| Inter-chiplet distance | 800 μm |
| TSV technology | TSV middle, $\varnothing 10\mu\text{m}$, height 100 μm , pitch 40 μm |
| #TSV | 14 000 TSV (2 000 signals + 12 000 power supply) |
| Backside RDL | 10 μm width, 20 μm pitch |
| C4-bumps | $\varnothing 90\mu\text{m}$, pitch 200 μm , 4,600 bumps |
| Flipchip package | BGA 39 x 39, 40mm x 40mm, 10 layers |
| Balls | $\varnothing 500\mu\text{m}$, pitch 1mm, 1 517 balls |

onto a package organic substrate (ten layers), using C4 bumps with a pitch of 200 μm .

In terms of complexity: 150 000 3D connections are performed using μ -bumps between the chiplets and the active interposer, with 20 000 connections for system communication, using the various 3D-Plugs, and 120 000 connections for power supplies using the SCVRs; while 14 000 TSVs are implemented for power supplies and off-chip communication. Due to the high level of complexity of the system, 3D assembly sign-off has been performed using the Mentor 3DStack CAD tool [50].

In Fig. 6, we present the respective floorplans of the chiplets and the active interposer. For the chiplet, one can see the four computing clusters and associated L1/L2/L3 caches, while for the active interposer one can see the different SCVRs, which are supplying power to each individual chiplet, the distributed system interconnects, and the system IOs on the circuit periphery. Dense 3D connectivity is done in various locations of the circuit using the 3D-plug interfaces and associated μ -bumps. Finally, the overall circuit has been packaged in a ball grid

array (BGA) flip-chip package with ten layers. In addition, one can see the six chiplets onto the package, before the final assembly with the cover lead and the package serigraphy.

More details on the various 3D technology element (μ -bumps, TSVs, RDL, and so on) and 3D assembly steps, with in-depth technology characterization, can be found in [23].

IV. COMPUTING CHIPLET ARCHITECTURE

A. Chiplet Overview

The focus of this architecture is its **scalability**, so we chose to design homogeneous chiplets embedding both processors and caches [24]. With the current memory mapping, the architecture can be tiled up to 8 \times 7 chiplets, last 2D-mesh row being reserved for off-chip IO accesses, achieving a maximum number of 56 chiplets, for a hypothetical total of 896 cores. The last-level cache is large enough with respect to computing power to release the pressure on the external memory access.

Each chiplet is composed of four clusters of four 32-bit scalar cores (MIPS32v1 compatible ISA) as shown in Fig. 7. System interconnects are extended to 3D using synchronous and asynchronous so-called “3D Plugs.” Chiplets form a single fully cache-coherent architecture composed of the following elements: separate 16-kB L1 Instruction-cache (I-cache) and Data-cache (D-cache) per core with virtual memory support, a shared distributed L2-cache with 256 kB per cluster, and an adaptive distributed L3-cache, with four L3 tiles (4 \times 1 MB) per chiplet.

All clocks (Cluster, L3 tile, and interconnect) are generated by 11 frequency locked loop (FLL) clock generators. To mitigate PVT variation, particularly across the dies in a 3D stack, we implement a timing fault methodology for Fmax/Vmin tracking [30]. Finally, chiplets are tested using IEEE1687 IJTAG, compressed full-scan, memory built-in self-test (BIST), and boundary scan for 3D IOs test, to allow for KGD assembly, as explained in Section VIII.

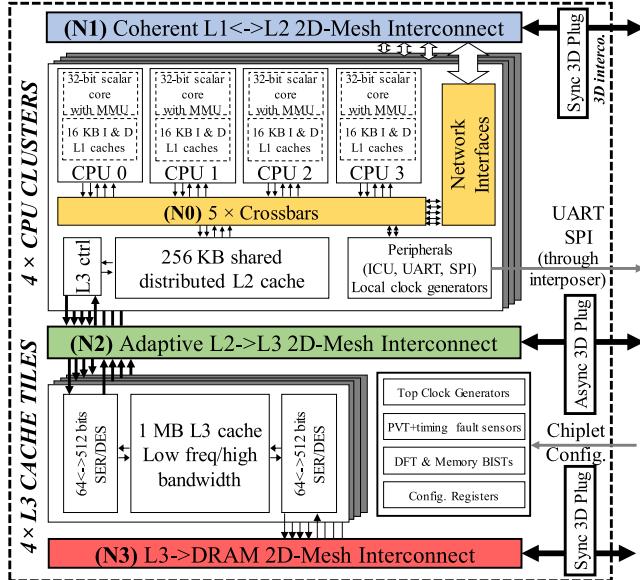


Fig. 7. Chiplet architecture, offering a 16-core scalable coherent fabric.

B. Chiplet Interconnects and Their Extension

Four different system level interconnects (N0 to N3) make up the system communication infrastructure as shown in Fig. 7, three of which are extendable in 3D: (N0) within cluster, an NoC crossbar allows the communication between the four cores through I& D caches and network interface; (N1) between L1 and L2 caches, a 5-channel 2D-mesh interconnect implements the coherency protocol and is routed in the interposer through passive links (two on each side); (N2) between L2 and L3 tiles, a 2-channel 2D/3D-mesh interconnect; (N3) between L3-caches and off-chip DRAM memory, a 2-channel 2D/3D-mesh interconnect. (N1) 2D-mesh is fully extended to other chiplets for maximum throughput and lowest short-reach latency as shown in Section VII. Peripherals are also connected to this network, which conveys IO requests. (N2) and (N3) networks implement a hierarchical routing scheme where a single router among the four of the chiplet 2D-mesh is used to reach the active interposer. This architecture reduces the 3D footprint for N2 and N3 networks, which are less bandwidth demanding. Using asynchronous logic for N2 3D plug allows for low latency L2 to L3 communications.

C. System Memory Mapping

The memory hierarchy is distributed and adaptive (Fig. 8): the 1-TB memory space is physically distributed among L2-caches accessed through (N1) network. Cluster coordinates in the 2D-mesh are encoded in the eight most significant bits of the address, forming a nonuniform memory architecture (NUMA), as done in [48]. Due to the X-first routing scheme of (N1) network, access to IO controllers located in the external FPGA is done through the North port of the ($X = 3, Y = 5$) router found in upmost right ($X = 1, Y = 2$) chiplet. Thus these IOs are mapped at [0 × 3 600 000 000: +40 GB] memory segment.

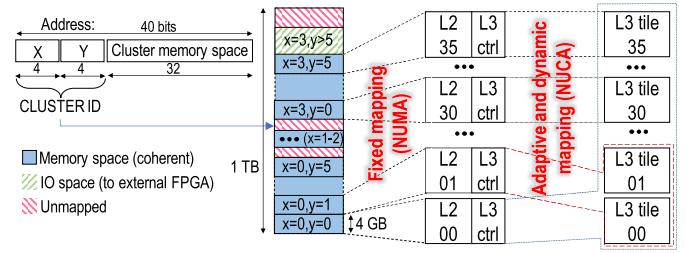


Fig. 8. Memory mapping and cache allocation.

D. Cache Coherency Scalability

1) **L1 Caches:** Each core has a private L1 cache that implements a Harvard's architecture with separate four-way 16-kB cache memories for instruction (I) and data (D) with 64-B cache lines. L1 D-caches are write-through and implement a fully associative write buffer composed of 8128-B entries flushed either on explicit synchronization or on expiration of an 8-cycle timer. L1 I/D-caches include a Memory Management Unit (MMU), which consists of two, per-core private, fully associative, 16-entry translation lookaside buffers (TLBs) for instruction and data. MMUs with coherent TLBs translate the 32-bit virtual address space (4 GB) in processor cores onto the 40-bit physical address space (1 TB) mapped as shown in Fig. 8. The hardware guarantees the coherency of both L1 I/D-caches and both I/D-TLBs (see Section IV).

As mentioned previously, the processor implements an NUMA memory hierarchy, and the most significant bits of the physical address designate the coordinates of the L2 cache containing that data. To improve performance, the operating system (OS) needs to map data and code close to the cores that use them. OS can do that through the virtual memory subsystem by the mapping of physical memory pages. To assist the OS in this task, our architecture implements two hint bits in the page table: the local (L) and remote (R) bits. They are automatically set by the MMUs and signal if a given physical memory page has been accessed by a core in the same (local) or different (remote) cluster than the L2 cache that hosts the cache lines in that page, also called the “home node.” For instance, pages with the R bit set but the L bit unset are candidates for migration.

2) **L2 Caches:** L2 caches are 16-way 256-kB set associative write-back caches handling 64-B cache lines. The scalable cache coherence protocol exploits the fact that shared data are mostly either sparsely shared read-write or widely shared read-only [25], [27]. Thus, L2 caches maintain L1-caches, TLBs, and IO coherence using a directory-based hybrid cache coherence protocol (DHCCP) based on write-through L1 caches. L2-cache lines have two states: a list mode where coherence is maintained by issuing multicast updates/invalidates; a counter mode where only broadcast invalidates are sent for this line.

In list mode, the sharers' set of this line is stored as a linked list: the first sharer ID is in the L2 directory and the following in another memory bank (heap). When a core writes to this line, the respective home L2 cache sends update messages to sharers, thus keeping their copy up to date.

When the number of sharers reaches a predefined threshold (4 in our implementation) or if the heap is full (4096 entries in our implementation), the cache line is put in counter mode where only the sharers' count is stored and broadcast invalidates are issued. The (N1) 2D-mesh and (N0) crossbar NoCs provide hardware support for broadcast and only L1 sharers of this line answer the broadcast, thus limiting the impact of broadcasts on scalability.

This hybrid sharing set representation is efficiently handling both main categories of shared data [26]. Write-through associated update messages also mitigate false sharing problems. The L2-cache coherence directory represents only 2% of die area with 15-bits core/cache IDs, showing the scalability of the cache coherence protocol. Section X-C shows scalability results for up to 512 cores.

3) L3 Caches: L3-cache tiles are 16-way 1-MB set associative write-back caches handling 128-B cache lines with one dirty bit per 64-B block. Tiles are dynamically allocated to L2-caches by software, forming a nonuniform cache architecture (NUCA) as presented in Fig. 8. In the case of L3 cache overlap, a shared tile behaves as a shared cache: more space is allocated to the most demanding memory segment. By overlapping L3 caches, the L3 cache controller located at the output of each L2 cache offers an L3 fault-tolerant adaptive repair. The controller uses a list of defective tiles to redirect traffic initially directed at these tiles to known working tiles. More detail on L3 micro-architecture and performance can be found in [28] and [29].

V. INTEGRATED VOLTAGE REGULATOR (VR)

A. Principle and 3D-Staking

Granular power delivery is a key feature to improve the overall energy efficiency of multi-core processors [31]. To allow DVFS per-chiplet, fast transitions, and mitigate IR-drop effects, six integrated VRs have been included in the interposer layer which individually supplies each chiplet by V_{core} from V_{in} as shown in Fig. 9. The power is delivered through the μ -bump flip-chip matrix. The V_{in} voltage is delivered from the interposer back-face through a 40- μ m-pitch TSV array. The VRs are fully integrated into the active interposer without needing any external component.

The typical input voltage range V_{in} is 1.5–2.5 V to reduce the delivered current I_{in} from TSV, package, and motherboard. Thus, the number of package's power IOs can be reduced compared to a direct V_{core} distribution from external VR. The power distributed network loss is also reduced.

B. Circuit Design

The SCVR has been chosen thanks to their fully integration capability [31]–[35]. The chosen topology is a parallel-series three-stage gearbox scheme to cover a large V_{out} range while maintaining power efficiency (Fig. 10). Thus, the SCVR generates seven lossless voltage-ratio from 4:1 to 4:3. From 1.8 V_{in} , the SCVR provides from 0.35 to 1.35 V, which covers the low-to-high chiplet's power modes. The gearbox scheme is interleaved into ten phases to reduce the V_{core} ripple and to increase the control bandwidth. The number

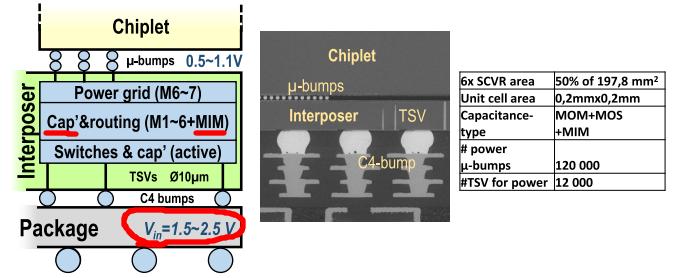


Fig. 9. SCVR cross section.

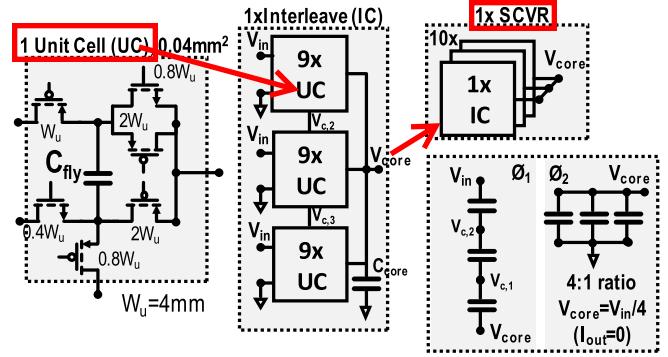


Fig. 10. SCVR unit-cell schematics and hierarchy.

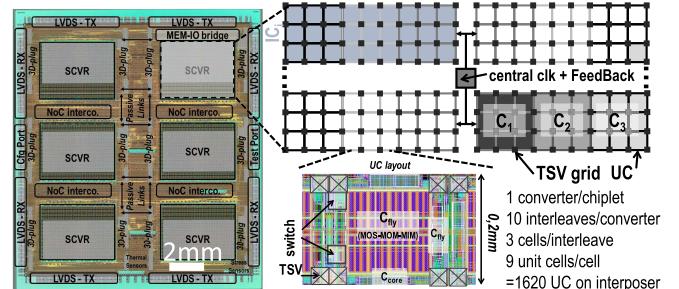


Fig. 11. SCVR layout on the interposer.

of interleaved phases is also chosen to maintain power efficiency at low-voltage level where the required power for chiplet drops off. The feedback control is based on one-cycle hysteresis controller proposed in [34]. The voltage controller is centralized and sequences the charge injection in the interleaved converters at each clock cycle. The clock generation and controller is integrated on-chip.

C. Physical Design on Interposer

As shown in Fig. 11, each SCVR occupies 50% of the chiplet footprint (11.4 mm^2) and is composed of 270 regular unit cells (UCs), with a 0, 2-mm pitch, in a checkerboard pattern. The I/O device transistor may operate on an up to 3.3-V input voltage. A MOS–MOM–MIM capacitor stack maximizes the capacitance density (8.9 nF/mm^2) with 102-nF flying capacitor per SCVR. To deal with potential process defaults on the large area of the interposer, fault-tolerant protocol is also included to mitigate the effect of defective unit cells on overall power efficiency.

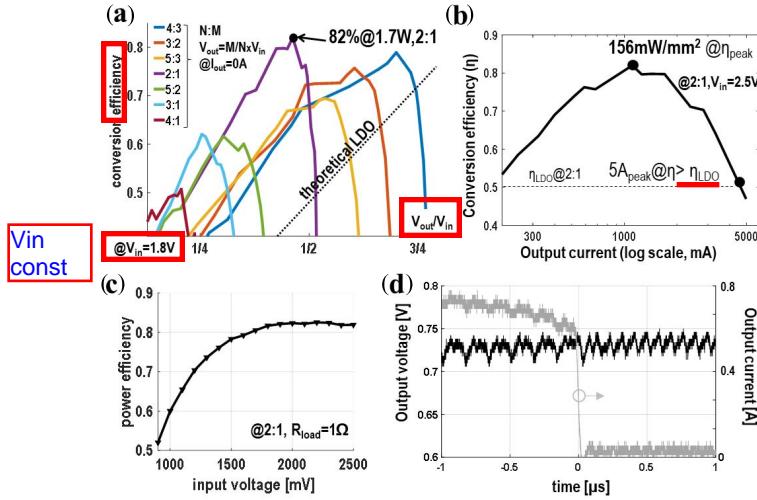


Fig. 12. SCVR experimental results. (a) Power efficiency versus voltage conversion ratio and gearbox configurations. (b) Efficiency over output current. (c) Efficiency versus input voltage at 2:1 ratio. (d) Load transient.

TABLE II

COMPARISON WITH COMPARABLE SCVR USING MOS OR MIM CAPACITOR TECHNOLOGY

| Reference | [32] | [33] | [35] | This work | Unit |
|---------------------|---------|------------|-------------|----------------|--------------------|
| Integration context | 2D | 2D | 2D | 3D | - |
| CMOS tech. | 65 | 90 | 28 | 65 | nm |
| Capacitor tech. | MOS | MOS | MIM | MOS+MIM | - |
| Ratio | 2 | 1 | 4 | 7 | - |
| Interleaving | 8 | 21 | 8 | 10 | - |
| Module area | 1.11 | 2.14 | 0.46 | 11.4 | mm ² |
| Flying cap density | 5.5 | 5.6 | 11.7 | 8.9 | nF/mm ² |
| V_{in} range | 2.3 | 2.3~2.6 | 1.8 | 0.9~2.9 | V |
| V_{out} range | 0.8~1.2 | 1.0 | 0.2~1.1 | 0.4~1.8 | V |
| Max power | 0.67 | 0.12 | - | 2.6 | W |
| Step response | 250 | 15,000 | 200 | 10 | ns |
| Droop voltage | 150 | 95 | 100 | 20 | mV |
| Peak efficiency | 71 | 69 | 73 | 81 | % |
| Power density | 550 | 770 | 310 | 156 | mW/mm ² |

D. Experimental Results

As shown in Fig. 12(a), the SCVR achieves a measured 156- and 351-mW/mm² power density at 82% peak efficiency and similar LDO's efficiency, respectively. SCVR maintains more than 60% from 0.45 to 1.35 V covering the full supply voltage range of the chiplet [Fig. 12(c)]. The SCVR delivers up to 5-A output current while maintaining higher efficiency than an LDO [Fig. 12(b)]. The peak power efficiency is relatively constant against V_{in} typical range. As shown in Fig. 12(d), the feedback control achieves less than 10-ns step response for a middle-to-zero load transient (0.8 to 0 A), while the full load is defined at peak efficiency (1.2 A).

Table II compares the 3D stacked SCVR to some previously published SCVR in 2D context. The proposed SCVR exhibits the highest number of lossless ratio and the highest delivered power with a commonly available capacitor to enable widely spread use. Even if the SCVR is affected by TSV grid, the power density is comparable to other wide range SCVRs. 3D integration of the SCVR on the interposer minimizes the system area and cost, with no impact on the chiplets.

TABLE III
3D-PLUG TYPES AND USAGE IN INTACT

| System Level Interconnect | 3D-Plug type | Active interposer link type |
|---|----------------------|-----------------------------|
| $L1 \leftrightarrow L2$ ($N1$ NoC) | Synchronous version | Short Reach, passive link |
| $L2 \leftrightarrow L3$ ($N2$ NoC) | Asynchronous version | Long reach, active link |
| $L3 \leftrightarrow ExtMem$ ($N3$ NoC) | Synchronous version | Long reach, active link |

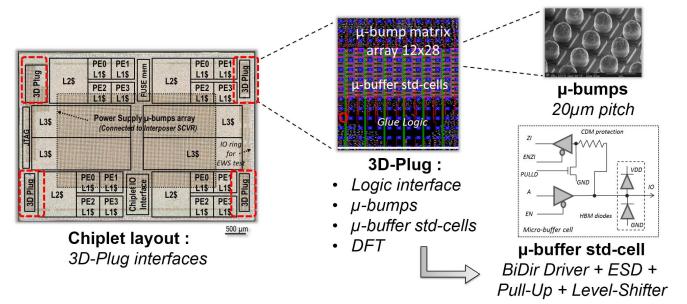


Fig. 13. 3D-Plug physical and logical interface overview.

E. Discussion

Since the power efficiency obtained by the integrated VR is lower than external dc-dc converters, the overall power efficiency of the computing system could improve by allowing fine-grain DVFS without increasing the bill-of-material (BoM) and IOs numbers. The power density is smaller than previously published results but the converters are fully integrated within the active interposer, not on the same die, thus reducing the cost impact of the proposed active power mesh. The interposer integration opens the opportunity for dedicated post-process high-density capacitors (e.g., deep trench capacitors) connected through TSV. We also prove the up-scaling capability of SCVR by fabricating the largest die area SCVR with a built-in capacitor fault-mitigation scheme.

VI. 3D-PLUG COMMUNICATION INTERFACES

A. 3D-Plug Features Overview

As presented in Section IV-B, the different chiplet system level interconnects are extended throughout the active interposer, by using generic chiplet–interposer interfaces, called 3D-Plugs.

Each 3D-Plug integrates both the logical and physical interfaces. As presented in Fig. 13, it contains the micro-bump array, the micro-buffer cells (bi-dir driver with ESD protection, level shifter, and pull-up), and boundary-scan logic for DFT. A bi-directional driver is used to allow testability of the interface before assembly (see Section VIII). This 3D interface is very similar to the 3D-NoC interface, as presented earlier in [36]. Due to the 28-/65-nm technology partitioning, the micro-buffer cell also requires in that case a level shifter in order to bridge the voltage domains between the chiplet (typically 1.0 V) and the active interposer (1.2 V).

In terms of physical design, the different 3D IOs of each 3D-plug have been created and assigned in an array fashion, while the micro-buffer cell has been designed as a standard cell and pre-placed within the pitch of the micro-bumps. All

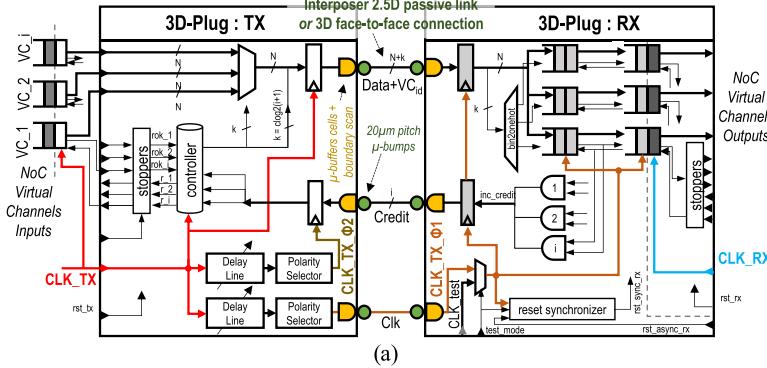


Fig. 14. (a) Synchronous 3D-Plug micro-architecture. (b) Comparison to state of the art.

the other parts of the 3D-Plug (their logical interface and DFT) have been designed using automated place and route.

In order to build the system level interconnects of INTACT, different kinds of 3D-plug have been designed, as presented in Table III.

Due to the different natures of the interconnects, in terms of traffic and distance/connectivity, two different kinds of 3D-Plugs have been designed, and compared in detail: one using synchronous design, as presented in Section VI-B, and one using asynchronous design, as presented in Section VI-C.

B. 3D-Plug Synchronous Version

The microarchitecture of the source synchronous 3D-Plugs used for 2.5D passive (N1 NoC) and 3D face-to-face links (N3 NoC) is shown in Fig. 14(a). Implemented as a standard synthesizable digital design, 3D-Plugs provide multiple virtual channels (VCs), the number of which is configured at design time. They use credit-based control flow and clock forwarding -Plug control logic operates at a higher frequency than the NoCs to reduce contention due to VC multiplexing. Delay lines and polarity selectors are used to skew TX clock for RX data sampling (CLK_TX_Φ1) and TX credit sampling (CLK_TX_Φ2).

When attached to the 3D vertical active link, the 3D-Plug achieves 3-Tb/s/mm² bandwidth density, 1.9× higher than [5]. 2.5D passive links reach a 12% higher bandwidth cross section than [5] as shown in Fig. 14(b). The aggregate synchronous 3D/2.5D links bandwidth is 527 GB/s.

We performed a frequency, logic voltage, and clock phase sweep on synchronous 2.5D/3D links. All 2.5D passive links were able to reach at least 1.25 Gb/s/pin in the [0.95 V-1.2 V] VDD range and the best link shown in Fig. 15 was able to reach this bandwidth at 1 V, while reaching more than 1.6 Gb/s/pin at 1.2 V. We obtained best results with a 180° CLK_TX_Φ1 phase and varying CLK_TX_Φ2 phase depending on frequency. While much shorter than passive links, 3D vertical links achieve slightly lower data rates of 1.21 Gb/s/pin upward and 1.23 Gb/s/pin downward as one side of these 3D-Plugs is implemented in the more mature and slower 65-nm technology of the interposer.

C. 3D-Plug Asynchronous Version

For its inherent robustness to any source of timing variations and low latency [37], asynchronous logic is well

| | This work | [5] VLSI'19 | Units |
|------------------------------------|--|--|----------------------|
| Technology | 28nm FDSOI chiplets / 65nm active interposer | 7nm FinFet chiplet | |
| 3D Link type and technology | 2.5D Passive (65nm Si) | 3D LIPINCON™ Active (face-to-face) | |
| System Integration | L1\$ ↔ L2\$ | L3\$ ↔ Ext. Mem | |
| Bus Width | 168 | 156 | 320 pins |
| # On-Chip Bus Links | 6+8+1 | 6 | 2 |
| Channel Length | 1.5 - 1.8 | 0.05 | mm |
| Die-to-Die Bump Pitch | 20 | 40 | μm |
| Voltage swing | 1.2 | 0.3 | V |
| Data Rate | 1.25 | 1.21 | Gb/s/pin |
| Power Efficiency | 0.75 | 0.59 | pJ/bit |
| Bandwidth Density | 0.5 | 3.0 | Tb/s/mm ² |
| Bandwidth Cross Section | 0.9 | N/A (unconstrained) | 1.6 Tb/s/mm |
| Aggregate inter-Chiplets Bandwidth | 394 | 133 | - |
| | 527 | | 640 GB/s |

(b)

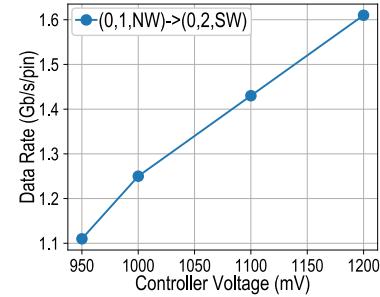


Fig. 15. Synchronous 3D-Plug max data rate for 2.5D passive links.

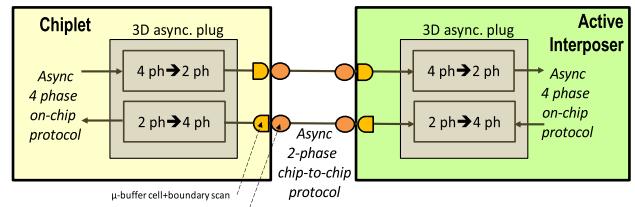


Fig. 16. 3D Plug asynchronous version overview, composed of protocol converters between the on-chip communication and the 3D interface.

adapted for designing system level interconnects and NoC architectures in a globally asynchronous locally synchronous (GALS) scheme. In the context of 3D architectures, asynchronous logic and its local handshakes enable interfacing two different dies without any clocking issues. By using robust quasi-delay insensitive (QDI) logic, an asynchronous 3D NoC has been earlier presented in [36] but presents some 3D throughput limitations due to the four-phase handshake protocol.

For INTACT, an innovative 3D-Plug interface has been designed, to benefit from two-phase handshake protocol at the 3D interface, which reduces the penalty of 3D interface delay within the interface cycle time, and thus increases the 3D interface throughput.

As introduced in [38], the principle is as follows (Fig. 16).

- Use asynchronous two-phase protocol for 3D interface communication, to reduce 3D interface delay penalty;
- Use asynchronous four-phase protocol for on-chip communication, within the active interposer, for its inherent simplicity, low latency and performance [37];

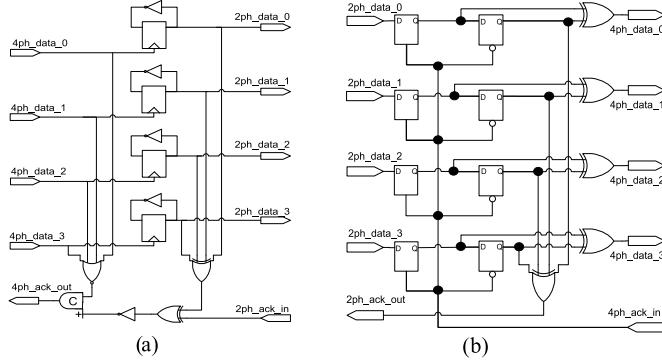


Fig. 17. 3D-Plug asynchronous version details, composed of (a) four-phase to two-phase protocol converter and (b) two-phase to four-phase protocol converter.

- Introduce a protocol converter, from two-phase protocol to four-phase protocol and respectively, using an *ad hoc* asynchronous logic encoding.

A recent overview of asynchronous logic and signaling can be found in [39]. For implementing a low cost protocol converter, a two-phase 1T-of-N multi-rail transition based signaling is used [38], with $N = 4$ (4-rail encoded, thus four wires for two bits). In this encoding and two-phase protocol, one single transition on Rail_i indicates the i value, which is then acknowledged by a transition on the feedback path. This encoding is close to the 1-of-n on-chip protocol, which leads to the corresponding protocol converters, shown in Fig. 17.

Similar to the synchronous 3D-plug interface, the protocol converter also integrates all the 3D objects: micro-bumps, micro-buffers, and boundary scan (Fig. 16). Finally, since the same number of wires is used for both protocols, a bypass mode of the protocol converters is added, configuring the 3D-plug interface either in two-phase mode or in four-phase mode, for circuit measurements.

VII. ACTIVE INTERPOSER SYSTEM INTERCONNECTS

A. Overview

Different kinds of system interconnects have been implemented between the chiplets on the interposer, using the 3D plugs described in Section VI. These interconnects are used to transport the different levels of cache coherence in the memory hierarchy. As discussed earlier, a mix of synchronous and asynchronous implementations was used, depending on latency and power targets.

The structure of the different interconnects is shown in Fig. 18, with clock-domain crossings, conversion interfaces, pipelining, and routers. These three interconnects will be detailed in the next paragraphs. To assess their performance, on-chip traffic generators and probes were inserted in the chiplets NoCs, for throughput and latency measurements.

B. L1–L2 Cache Interconnect

Fig. 18(a) presents the first level of cache interconnect between local L1 and distributed L2 caches (N1 NoC). As this first level of cache traffic is intended to be localized using an adequate application mapping, most of the traffic is expected to be exchanged between neighboring chiplets. Aside from clock-domain crossing between the two chiplets using

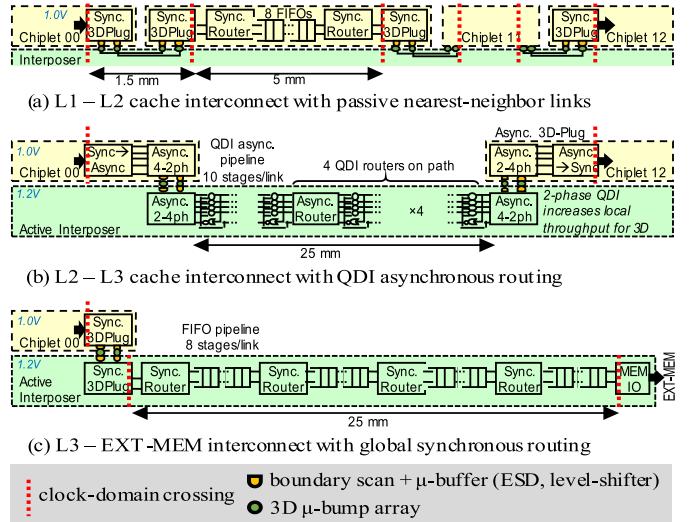


Fig. 18. INTACT system interconnect structure on longest path using different technologies for different traffic classes. (a) L1 and L2 cache interconnect with passive nearest-neighbor links. (b) L2 and L3 cache interconnect with QDI asynchronous routing. (c) L3–EXT-MEM interconnect with global synchronous routing.

synchronous 3D-Plugs, no other adaptation is required, and routing is entirely performed within the chiplets. Therefore, only passive metal wires are used on the interposer to connect the microbumps of neighboring chiplets.

Physical design of these interposer passive links was optimized to reduce delay and crosstalk between the nets. A dedicated routing scheme on two levels of metal was used (M3–M5 horizontal, M2–M4 vertical), with trace widths of 300 nm and spacing of 1.1 μ m. Additional shielding was used for clock nets running at twice the data rate. Crossings with minimum-width unrelated wires on the interposer showed very little impact on crosstalk or delays in the signal, and were therefore allowed on the other interposer metal levels.

Point-to-point connection between two adjacent 3D-Plugs was measured at 1.25 GHz, with a latency of 7.2 ns. Most of this latency is due to the clock-domain crossings in 3D-Plugs.

For large applications, nevertheless, L1–L2 cache coherence traffic needs to extend farther than between adjacent chiplets. In that case, pipelining and routing are handled by the intermediate chiplets. The main advantage in this case is that this is done using the advanced technology node in the chiplets, which has a better performance and lower power consumption than the interposer does. However, the major drawback is the accumulation of pipeline and clock-domain crossings, which adds extra latency for distant L1–L2 traffic.

The 2D NoC frequency in the chiplet runs at 1 GHz, but the one-way latency from the source 3D-Plug to the destination 3D-Plug can be as high as 44 cycles on the longest path from chiplet 00 to chiplet 12, with two intermediate chiplets, five routers and eight to ten FIFO stages between routers. Nevertheless, this solution is very energy efficient with only 0.15 pJ/bit/mm.

C. L2–L3 Cache Interconnect

Fig. 18(b) presents the second level of cache interconnect between distributed L2 and L3 caches (N2 NoC). The main

TABLE IV
COMPARATIVE PERFORMANCE OF SYSTEM INTERCONNECTS IN INTACT

| | L1-L2 nearest | L1-L2 farthest | L2-L3 4-phase | L2-L3 2-phase | L3-EXT-MEM | 3DNOC [36] | Units |
|--------------------|----------------|----------------------|---|-----------------------|----------------------|-----------------------|-----------|
| Reach | 1.5 | 15 | 25 (bottom left chiplet to upper right chiplet) | | 8 | | mm |
| Word size | | 40 | | 72 | | 32 | bits |
| Interposer | 1 passive link | 3 passive links | Active async. routing | Active async. routing | Active sync. routing | Active async. Routing | — |
| Chiplet | — | Global sync. routing | Local sync. routing | Local sync. routing | Local sync. routing | — | — |
| 3D Plug frequency | 1.25 | 1.25 | 0.30 | 0.52 | 1.21 | 0.32 | GHz |
| 2D NoC frequency | — | 1.00 | | 0.97 | 0.75 | 0.89 | GHz |
| End to end latency | 2x4+[0-1] | 44 | 4 + async. | 4 + async. | 37 | 4 + async | Cycles |
| | 7.2 | 44.0 | 15.2 | 15.2 | 49.5 | 10 | ns |
| Propagation speed | 4.8 | 2.9 | 0.6 | 0.6 | 2.0 | 1.2 | ns/mm |
| Energy / bit / mm | 0.29 | 0.15 | 0.52 | 0.52 | 0.24 | 0.5 | pJ/bit/mm |

performance target is in this case to offer low latency long reach communication. For this purpose, it was chosen to implement it in fully asynchronous logic on the interposer, using the ANoC QDI NoC [37]. This allows for only two synchronous/asynchronous conversions on an end-to-end path, to save on clock-domain-crossing latency. Deep-pipelining on the ANoC allows to insert an asynchronous pipeline stage every 500 μ m to preserve throughput with almost no impact on the latency compared to inverter-based buffering.

The asynchronous 3D-Plug in two-phase mode allows an injection rate in the network for 72-bit data words up to 520 MHz, while the 2D NoC is able to sustain up to 0.97 GHz on every link, which limits the in-network contention of overlapping network paths. The efficient asynchronous pipelining allows an end-to-end latency on the synchronous interfaces of the 3D-Plugs of only 15.2 ns, with four clock cycles and 11.2 ns of asynchronous latency across four routers and 25 mm of pipelined links.

D. L3-ExtMemory Interconnect

Fig. 18(c) presents the last interconnect between the distributed L3 caches and the external memory (N3 NoC). Considering the intrinsic contention of this last level of cache traffic, and the longer latency for paginated access to the external memory, the focus was put on energy efficiency, then on low latency. This interconnect is implemented as a global synchronous NoC, with clock-domain crossings at the source 3D-Plug and in the memory IO interface. Two-stage FIFOs are inserted every 1 mm, and tight clock-tree balancing was performed to increase the throughput. This results in a 72-bit synchronous network running up to 750 MHz, with a latency of 2 ns/mm, for a good energy efficiency of 0.24 pJ/bit/mm.

E. System Interconnect Comparison and Conclusion

Table IV summarizes the different figures of merit for the three interconnects, and provides a benchmark with respect to the 3D NoC in [36]. It shows that neighboring connections can be efficiently made using the synchronous 3D-Plug in an advanced technology node, with a high throughput and a low power consumption. For longer-range communication, limiting the number of clock-domain crossings is key for performance. The NoCs in the active interposer can provide

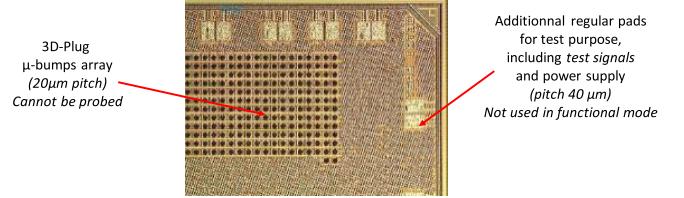


Fig. 19. Chiplet layout (zoom), with 3D-Plug interface and additional test pads.

wide interconnects optimized for latency in the asynchronous version, with 0.6 ns/mm, or for power consumption in the synchronous version, with 0.24 pJ/bit/mm, with performance metrics twice as good as [36] in the same 65-nm technology node as the active interposer.

The achieved low level interconnect performances could be used for a more systematic system level study, such as [19], by trading off different traffic classes, latency, and energy, thanks to the extended active interposer traffic capabilities.

VIII. ACTIVE INTERPOSER TESTABILITY AND 3D DFT ARCHITECTURE

A. Testability Challenges

With such 3D active interposer, testability is raising various challenges. First, it is required to ensure KGD sorting to achieve high system yield [10]. This implies that the 3D test architecture must enable EWS test of the chiplet and the interposer (pre-bond test, before 3D assembly), and final test (post-bond, after 3D assembly in the circuit package). Moreover, due to fine pitch μ -bumps, reduced test access is observed, μ -bumps cannot be directly probed in test mode. This implies to include additional IO pads, which are only used for test purpose, and not in functional mode (see Fig. 19).

Finally, with 3D technologies, additional defects may be encountered, such as μ -bumps misalignments, TSV pinhole, shorts, and so on which lead to specific care for testing the 3D objects and interfaces. Another concern is also regarding the automatic test pattern generation (ATPG) engineering effort, where easy re-targeting of test patterns from pre-bond test to post-bond test should be proposed to reduce test development efforts.

Numerous researchers have addressed specific test solutions for 3D defaults, see for instance [40], [41], for testing generic

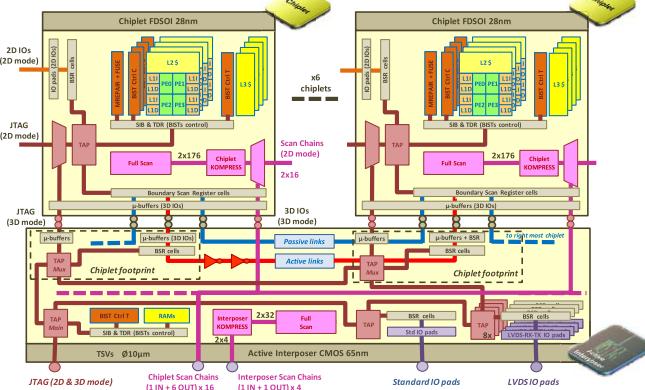


Fig. 20. 3D design-for-test architecture for INTACT, overview and detailed.

3D architectures using die wrappers and elevators [42], and for testing 2.5D passive interposers [43]. A standardization initiative on 3D testability has emerged with the recent IEEE 1838 standard [44]. Nevertheless, no work addressed initially the testability of active interposers.

B. 3D DFT Architecture

Within the INTACT architecture, the test of the 3D system must address the test of all the following elements: 1) the regular standard cell-based logic; 2) all memories using BIST engines and repair; 3) the distributed 3D interconnects and IOs: 3D connections of active links and passive links, which are implemented by micro-bumps; and finally 4) the regular package IO pads for off-chip communication through the TSVs.

In order to test the active interposer and its associated chiplets, the proposed 3D DFT architecture (Fig. 20) is based on the two following main test access mechanisms (TAMs), as proposed earlier in [45].

- A IJTAG IEEE1687 hierarchical and configurable chain, accessed by a primary JTAG TAP port, for testing all the interconnects and memories, based on the concept of “chiplet footprint.”
- A Full Scan logic network using compression logic, for reduction of test time and of number of test IOs.

By using IJTAG IEEE 1687, the JTAG chain is hierarchical and fully configurable: the JTAG chain provides dynamic access to any embedded test engines. The active interposer JTAG chain is designed similar to a chain of TAPs on a PCB board. It is composed of “chiplet footprints,” which provide either access to the above 3D-stacked chiplet or to the next chiplet interface, and which are chained serially. The JTAG network is used to test and control the 3D active links, the 3D passive links, the off-chip interfaces, and the embedded test engines, such as the memory BISTS.

The Full Scan logic network offers efficient and parallel full scan test of the whole 3D system logic. In order to reduce the number of 3D parallel ports, compression logic is used in both the chiplets and the active interposer, with a classical tradeoff (shift time/pin count). Independent scan paths are used

TABLE V
INTACT DFT RESULTS

| DFT access | Active Interposer 65nm | Chiplet FDSOI 28nm | Full 3D System |
|---|--|--|---|
| Full Scan | 32 scan chains, 4 after compression, #faults 5.7M, Test cov. ~60%**, 1318 patterns | 182 scan chains, 16 after compression, #faults 21.5M, Test cov. 97.1%, 1790 patterns | #faults 134.8M Test cov. 95.5% |
| IJTAG + Interco. Boundary Scan | All IO pads <i>pre-bond</i> 2D pads (826) 3D IO (13 548) 81 patterns | All IO pads <i>pre-bond</i> 2D pads (249) 3D IO (2 258) 68 patterns | 3D IO pads <i>post-bond</i> (13 548) 27 patterns |
| BIST & Repair | #BIST: 1 12 patterns / BIST | #BIST: 5 20 patterns / BIST | #BIST: 31 612 patterns |

** Limited test coverage is reported by the tool within the interposer, this is due to the asynchronous NoC that can be tested using a dedicated test solution not reported here

between the chiplets and the active interposer, to facilitate the test architecture integration.

C. Test CAD Flow and Test Coverage

The proposed 3D DFT architecture has been designed and inserted using Tessent tools from Mentor, a Siemens Business, Montbonnot, France. By using IJTAG and IEEE1687, high-level languages such as “Instrument Connectivity Language” (ICL) and “Procedural Description Language” (PDL) are provided and enable to handle the complexity of such a system. In particular, it is possible to fully automate the test pattern generation of Memory BIST engines, from ATPG at chiplet level to ATPG of the same patterns within the full 3D system, enabling so-called test pattern retargeting. As presented in Table V, full testability is achieved for all logic, 3D interconnects and regular package IOs, and memory BIST engines, before 3D assembly and after 3D assembly.

Using the proposed DFT architecture and test patterns, the full system was tested using an automated test equipment (ATE).

- The 28-nm chiplet has been tested at wafer level using a dedicated probe card, with a binning strategy.
- The active interposer has not been tested at wafer level, supposing the maturity of the 65-nm technology and its high yield due to its low complexity (see Section III-B). Nevertheless, its standalone DFT and dedicated IOs were initially planned and designed as mentioned above.
- The full INTACT circuit, after 3D assembly and packaging, has been tested within a dedicated package socket.

IX. THERMAL CHALLENGES AND STRATEGY

A. Thermal Challenges

In 3D technology, thermal dissipation is a challenge that needs to be properly addressed. Due to more integration in a smaller volume, a larger power density is usually observed in 3D, while the thermal dissipation itself is getting more complex in the overall 3D stack of the circuit and package, overall leading to thermal hotspots or even thermal runaway [52]. In the generic context of logic-on-logic stacking, thermal dissipation is worse because multiple layers of compute dies

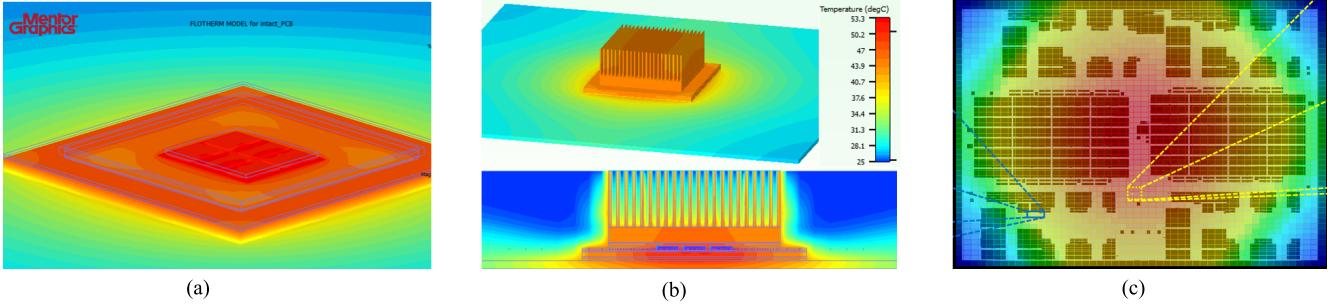


Fig. 23. (a) Package temperature (without heat sink). Peak temperature = ~150 °C. (b) Package temperature (with heat sink and fan). Peak temperature = 53 °C. (c) Chiplet thermal map. Peak temperature = ~53 °C.

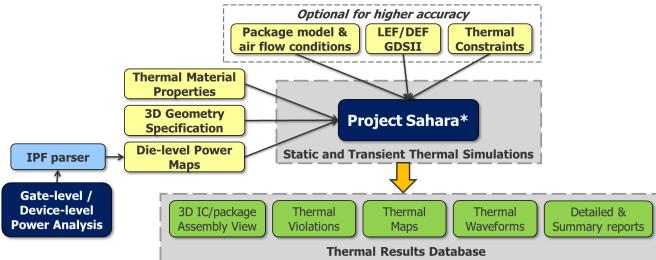


Fig. 21. 3D chip-package thermal flow, from early exploration to sign-off.

need to dissipate their heat on top of themselves. On the contrary, in the case of interposer-based systems, a single layer of chiplets is dissipating heat, while heat extraction can be performed from the top package face, similar to a regular flip-chip packaged circuit. Nevertheless, contrarily to passive interposers, in the case of an active interposer, the bottom layer is also part of the power budget, and dissipates heat as well. Since the power budget of the active interposer layer is rather limited, with most power budget within the chiplets, this should help the overall thermal dissipation.

Finally, due to the heterogeneous structure of such a 3D stack, many materials are composing the device, with silicon substrate, back-end of line (BEOL) in copper, underfill composite materials between the chiplets and interposer, micro-bumps (SnAg), TSVs (copper), and so on. This assembly leads to strong anisotropic thermal effects, favoring and increasing thermal hotspots effects. Moreover, due to the thin layer effect of the interposer (100 μm), the horizontal thermal dissipation is reduced in the interposer, while it remains mostly the vertical thermal dissipation through the chiplets. These various thermal effects have been widely studied in the literature [53], [54], and need to be taken into account in the full system.

B. Thermal Modeling Strategy

With all the 3D thermal challenges: increased power density and design complexity on the design side, fine grain material effects on the technology side, coupled to the regular package and board thermal information, an accurate thermal exploration must be performed with the adequate thermal methodology. Various thermal tools are available: either circuit

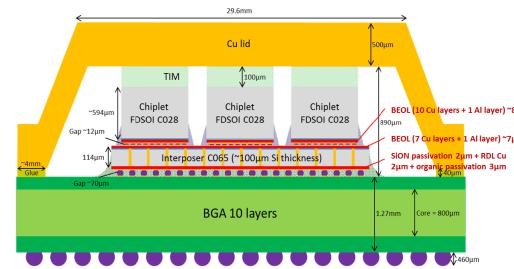


Fig. 22. INTACT circuit and package cross section used for thermal modeling.

level tools able to cope with detailed circuit and technology description but with simple packaging condition, or package level tools able to cope with detailed packaging, but with reduced die and technology information. In order to achieve an accurate thermal exploration covering all modeling aspects, the Project Sahara solution, a thermal analysis prototype from Mentor Graphics a Siemens Business, was selected [55].

As presented in Fig. 21, an adequate thermal methodology has been setup to allow modeling of low level structures (TSV, micro-bumps, underfill), with a design entry at GDS level and with accurate static or dynamic power maps, all this in the context of the full system (package and fan). The methodology has been qualified on a previous 3D logic-on-logic design with silicon thermal measurements 16 [36]. More details of the thermal methodology can be found in [56].

C. Thermal Simulation Results

The INTACT circuit and package have been modeled, as presented in Fig. 22 with a detailed cross section of the 3D circuit. In terms of power budget, a scenario with a maximum static power budget of 28 W is simulated, corresponding to a worst case situation of 3 W per chiplet ($\times 6$) and 10 Watts in the active interposer, while the nominal circuit power budget is 17 W as presented in Section X-B.

As a result, Fig. 23 shows the thermal exploration, without Heat Sink (max temperature 150 °C), with a regular Heat Sink and Fan (max temperature 53 °C), while no hotspots appear within the computing chiplet. Even for this worst case scenario, due to a still limited power density of 0.14 W/mm², the thermal dissipation of the active interposer can be achieved using a regular heatsink and fan.

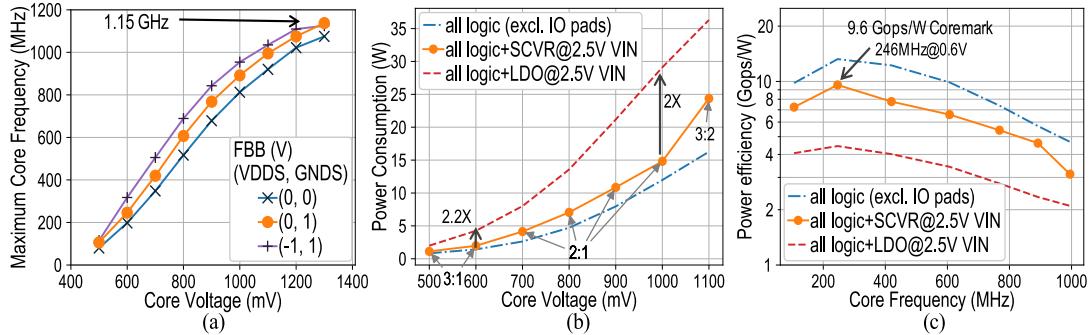


Fig. 25. (a) Maximum core frequency. (b) Power consumption at F_{\max} [$FBB = (0,1)$]. (c) Power efficiency at V_{\min} .

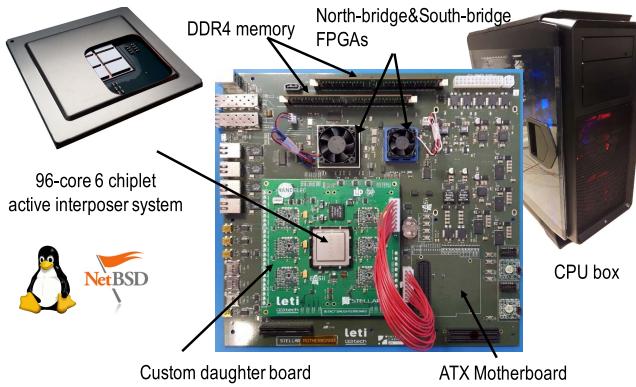


Fig. 24. Development board fitting in a standard PC case.

X. OVERALL CIRCUIT RESULTS

As shown in Fig. 24, a complete development board has been designed for measurement and application evaluations including running Linux on the chip. The board features two FPGAs with a 2×16 GB 64-bit DDR4 memory and various peripherals: 8x PCIe Gen3, SATA, 1-Gb Ethernet, 10-Gb Ethernet, HDMI, USB, SD-Card, and UART. The demonstration board also features a power infrastructure with voltage and current sensing. Each FPGA is connected to two of the four LVDS links of the chip.

A. Circuit Performances

The chiplet is functional in the 0.5–1.3-V range with Forward Body Biasing (FBB) [46] up to ± 2 V. Fig. 25(a) shows that a core frequency of 1.15 GHz is achieved at 1.3 V with 0/+1 (VDDS/GNDS) FBB. Single core performance is 2.47 Coremark/MHz and 1.23 DMIPS/MHz. At chip level, maximum energy efficiency is 9.6 GOPs/W on Coremark benchmarks (IPC = 0.8/core) at 0.6 V taking into account voltage regulation losses in the interposer as shown in Fig. 25(c). As expected, FBB boosts performance: in typical at 0.9 V, a frequency increase of 24% is achieved with -1/+1 FBB, while in typical at 680 MHz, an energy efficiency increase of 15% is achieved with asymmetric 0/+1 FBB.

B. Circuit Power Budget and Energy Efficiency

In Fig. 25(b) and (c), we show overall chip power and performance measurements with a 0/+1 FBB. Power

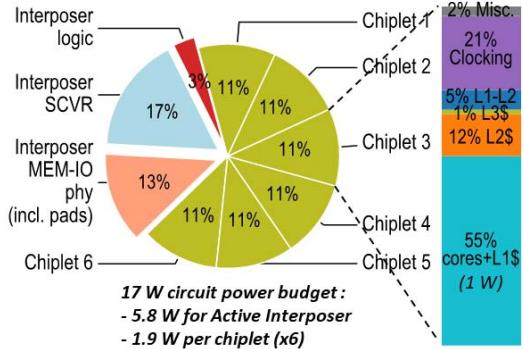


Fig. 26. Power consumption breakdown, cores operating at 1 V, 900 MHz.

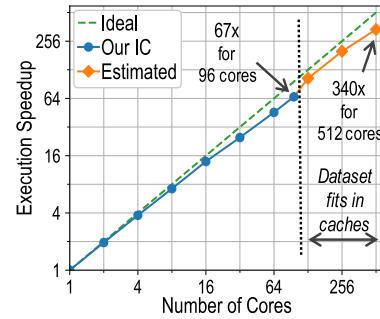


Fig. 27. Execution speedup up to 512 cores.

consumption and energy efficiency while running Coremark benchmark is compared to a theoretical system using a digital LDO instead of the proposed fully integrated SCVR. Using an LDO at the same $V_{\text{IN}} = 2.5$ V would result in a $2\times$ increase in power consumption, a lower V_{IN} would be needed to limit losses at the expense of more power pins and voltage-drop issues.

The power breakdown in Fig. 26 shows the low power budget of the active interposer with only 3% of total power consumed by the active interposer logic. The cores+L1\$ represent over half the power consumption of the chiplets, themselves consuming the majority of the measured circuit power (17 W).

C. Circuit Scalability

Lastly, Fig. 27 shows the scalability of the cache-coherent architecture that is analyzed by running a 4 Mpixels image

TABLE VI
STATE OF THE ART COMPARISON

| | <i>This work</i> | [31] ISSCC'18 | [4,15] ISSCC'18&20 | [5] VLSI'19 | [6] ISSCC'17 | <i>Units</i> |
|--------------|-----------------------------|--|-----------------------------------|----------------------------|-----------------|----------------------|
| | | INTEL | AMD | TSMC | INTEL | |
| Technology | Chiplet Technology | FDSOI 28nm | FinFET 14nm | FinFET 14nm/7nm | FinFET 7nm | FinFET 14nm |
| | Interposer Technology | Active CMOS 65nm | no | MCM substrate | Passive CoWoS ® | EMIB bridge |
| | Interposer extra features | yes | N/A | no / IO die | no | no |
| | Total system yield | high (mature tech. & low transistor count) | N/A | high | high | high |
| | Die-to-Die µbump pitch | 20 | N/A | > 100 | 40 | 55 <i>um</i> |
| Power | Voltage Regulator (VR) type | 6 SCVR on interposer with MOS+MOM+MIM | on-chip distributed SCVR with MIM | LDO per core, with MIM | no | no |
| | VR area | 34% of active interposer | MIM>40% core area | - | N/A | N/A |
| | VR peak efficiency | 82% | 72% | LDO limited | N/A | N/A |
| Interconnect | Interconnect types | Distributed scalable cache-coherent NoCs | N/A | Scalable Data Fabric (SDF) | LIPINCON™ links | AIB interconnect |
| | 3D Plug power efficiency | 0.59 | N/A | 2.0 | 0.56 | pJ/bit |
| | BW density | 3.0 | N/A | - | 1.6 | Tb/s/mm ² |
| | Aggregate 3D bandwidth | 527 | N/A | - | 640 | GByte/s |
| CPU | Number of chiplets | 6 | 1 | 1 - 4 / 1 - 8 | 2 | 1 FPGA + 6 TxRx |
| | Number of cores | 96 | 18 | 8 - 32 / 8 - 64 | 8 | FPGA fabric |
| | Max Frequency | 1.15 | 0.4 | 4.1 / 4.7 | 4 | 1 GHz |
| | Gops (32b-Integer) | 220 (peak mult./acc.) | 14.4 | 131.2 - 1203 | 128 | N/A |
| | | | | | | Gop/s |

filtering application from 1 to 512 cores. The filter is composed of a 1-D convolution, followed by a transposition of the image and ends with another 1-D convolution. Software synchronization barriers separate these steps and the transposition, in particular, involves many memory transfers.

Results for more than 96 cores were obtained by RTL simulation with additional chiplets. Software is executed on a single cluster up to four cores and on a single chiplet up to 16 cores. Compared to a single core execution, a 67× execution-time speedup is obtained with 96 cores and 340× with 512 cores. The slight uptick above 128 cores results from the threshold where the data set fits in caches. This quasi-linear speedup, ignoring limitations of the external memory bandwidth, shows the scalability of network protocols and their 3D implementations.

D. Comparison to Prior Art

Compared to prior art (Table VI), the INTACT circuit is the first CMOS active interposer validated on silicon, which offers a chiplet-based many-core architecture for HPC. The active interposer solution allows for integrated VRs without any external passives, using free die area available in the active interposer, offering DVFS-per-chiplet and achieving 156 mW/mm² at 82% peak power efficiency, with 10%–50% more efficiency with respect to LDO converters integrated in organic schemes. The SCVR is also fault tolerant to mitigate the effect of defective unit cells on the overall power efficiency.

Regarding interconnects, contrary to previous point-to-point solutions, the active interposer offers flexible and distributed NoC meshes enabling any chiplet-to-chiplet communication for scalable cache-coherency traffic, with 0.6-ns/mm inter-chiplet latency using asynchronous signaling within the interposer, and a 0.59-pJ/bit synchronous 3D-Plug energy efficiency with 3-Tb/s/mm² bandwidth density, which is twice better than previous circuits.

The overall system integrates a total of 96-cores, in six chiplets, offering a peak computing power of 220 GOPs (peak mult-acc), which is quite comparable to advanced state of the art processor systems. Finally, the overall distributed interconnects and cache coherency memory architecture are scalable up to 896 cores, showing the architecture partitioning capability to target larger computing scale.

XI. CONCLUSION

The presented *Active* interposer leverages the 3D integration benefits by offering a baseline of functionalities such as voltage delivery, chiplet-to-chiplet communications, IOs, shared by most of computing assemblies. The active interposer allows a flexible assembly with common functionalities while maintaining the yield management benefits. For this reduced power density and budget, thermal dissipation is not an issue within the active interposer, as for a regular passive interposer.

3D integration and active interposer open the way toward efficient integration of large-scale chiplet-based computing systems. Such scheme can be applied for integration of similar chiplets as presented in this article, but also for smooth integration of heterogeneous computing chiplets [47].

ACKNOWLEDGMENT

The authors would like to thank STMicroelectronics and Didier Campos team for INTACT package design and assembly, PRESTO Engineering and Brice Grisollet for testing the INTACT circuit onto Automatic Test Equipment, Easii-JC and Jean-Paul Goglio and his team for designing the INTACT application demonstration board. Finally, they would like to thank many other contributors from Mentor Graphics on the CAD tool teams and from CEA-LETI on both the design team and technology teams for their dedication to make this concept and circuit a successful realization.

REFERENCES

- [1] G. Agosta *et al.*, "Challenges in deeply heterogeneous high performance systems," in *Proc. 22nd Euromicro Conf. Digit. Syst. Des. (DSD)*, Aug. 2019, pp. 428–435.
- [2] P. Ramm, *Handbook of 3D Integration: Technology and Applications of 3D*, vol. 1. Hoboken, NJ, USA: Wiley, 2008.
- [3] M.-F. Chen, F.-C. Chen, W.-C. Chiou, and C. H. Doug Yu, "System on Integrated Chips (SoIC(TM) for 3D Heterogeneous Integration," in *Proc. IEEE 69th Electron. Compon. Technol. Conf. (ECTC)*, 2019, pp. 1–5.
- [4] N. Beck, S. White, M. Paraschou, and S. Naffziger, "Zeppelin": An SoC for multichip architectures," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2018, pp. 40–41.
- [5] M.-S. Lin *et al.*, "A 7nm 4GHz Arm-core-based CoWoS chiplet design for high performance computing," in *Proc. Symp. VLSI Circuits*, Jun. 2019, pp. 28–32.
- [6] D. Greenhill *et al.*, "A 14nm 1GHz FPGA with 2.5D transceiver integration," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2017, pp. 54–55.
- [7] P. Gupta and S. S. Iyer, "Goodbye, motherboard. Bare chiplets bonded to silicon will make computers smaller and more powerful: Hello, silicon-interconnect fabric," *IEEE Spectr.*, vol. 56, no. 10, pp. 28–33, Oct. 2019.
- [8] CHIPS Program. Accessed: 2017. [Online]. Available: <https://www.darpa.mil/program/common-heterogeneous-integration-and-ip-reuse-strategies>
- [9] 3 Ways Chiplets are Remaking Processors. Accessed: Apr. 2020. [Online]. Available: <https://spectrum.ieee.org/semiconductors-processors/3-ways-chiplets-are-remaking-processors>
- [10] J. Quinne and B. Loferer, "Quality in 3D assembly—Is, known good die good enough?" in *Proc. IEEE Int. 3D Syst. Integr. Conf. (3DIC)*, 3DIC'2013, pp. 1–5.
- [11] K. Sohn *et al.*, "18.2 a 1.2 V 20nm 307GB/s HBM DRAM with at-speed wafer-level I/O test scheme and adaptive refresh considering temperature distribution," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Jan. 2016, pp. 316–317.
- [12] T. F. Wu *et al.*, "14.3 a 43pJ/Cycle non-volatile microcontroller with 4.7s Shutdown/Wake-up integrating 2.3-bit/Cell resistive RAM and resilience techniques," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2019.
- [13] A. Jouve *et al.*, "Die-to-wafer direct hybrid bonding demonstration with high alignment accuracy and electrical yields," in *Proc. Int. 3D Syst. Integr. Conf. (3DIC)*, Oct. 2019, pp. 1–5.
- [14] Open Compute ODSA Project. Accessed: 2019. [Online]. Available: <https://www.opencompute.org/wiki/Server/ODSA>
- [15] S. Naffziger, K. Lepak, M. Paraschou, and M. Subramony, "2.2 AMD chiplet architecture for high-performance server and desktop products," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 44–45s
- [16] W. Gomes *et al.*, "8.1 lakefield and mobility compute: A 3D stacked 10nm and 22FFL hybrid processor system in 12×12 mm², 1 mm Package-on-Package," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 144–146.
- [17] G. Hellings *et al.*, "Active-lite interposer for 2.5 & 3D integration," in *Proc. Symp. VLSI Technol. Circuits*, 2015, pp. 222–223.
- [18] S. Chéramy *et al.*, "The active-interposer concept for high-performance chip-to-chip connections," *Chip Scale Rev.*, vol. 5, p. 35, Jun. 2014.
- [19] J. Yin *et al.*, "Modular routing design for chiplet-based systems," in *Proc. ACM/IEEE 45th Annu. Int. Symp. Comput. Archit. (ISCA)*, Jun. 2018, pp. 726–738.
- [20] V. Pano, R. Kuttappa, and B. Taskin, "3D NoCs with active interposer for multi-die systems," in *Proc. 13th IEEE/ACM Int. Symp. Networks-on-Chip*, Oct. 2019, pp. 1–8.
- [21] P. Vivet *et al.*, "2.3 a 220GOPS 96-core processor with 6 chiplets 3D-stacked on an active interposer offering 0.6ns/mm latency, 3Tb/s/mm² inter-chiplet interconnects and 156 mW/mm² 82%-Peak-Efficiency DC-DC converters," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 1–85.
- [22] D. Gitlin *et al.*, "Generalized cost model for 3D systems," in *Proc. IEEE SOI-3D-Subthreshold Microelectron. Technol. Unified Conf. (S3S)*, Oct. 2017, pp. 1–3.
- [23] P. Coudrain *et al.*, "Active interposer technology for chiplet-based advanced 3D system architectures," in *Proc. IEEE 69th Electron. Compon. Technol. Conf. (ECTC)*, May 2019, pp. 569–578.
- [24] E. Guthmuller *et al.*, "A 29 Gops/Watt 3D-ready 16-core computing fabric with scalable cache coherent architecture using distributed L2 and adaptive L3 caches," in *Proc. IEEE 44th Eur. Solid State Circuits Conf. (ESSCIRC)*, Sep. 2018, pp. 318–321.
- [25] J. Dumas, E. Guthmuller, and F. Petrot, "Dynamic coherent cluster: A scalable sharing set management approach," in *Proc. IEEE 29th Int. Conf. Appl.-Specific Syst., Archit. Processors (ASAP)*, Jul. 2018, pp. 1–8.
- [26] J. Dumas, E. Guthmuller, C. F. Tortolero, and F. Petrot, "Trace-driven exploration of sharing set management strategies for cache coherence in manycores," in *Proc. 15th IEEE Int. New Circuits Syst. Conf. (NEWCAS)*, Jun. 2017, pp. 77–80.
- [27] Y. Fu, T. M. Nguyen, and D. Wentzlaff, "Coherence domain restriction on large scale systems," in *Proc. 48th Int. Symp. Microarchitecture*, 2015, pp. 686–698.
- [28] E. Guthmuller, I. Miro-Panades, and A. Greiner, "Adaptive stackable 3D cache architecture for manycores," in *Proc. IEEE Comput. Soc. Annu. Symp.*, Aug. 2012, pp. 39–44.
- [29] E. Guthmuller, I. Miro-Panades, and A. Greiner, "Architectural exploration of a fine-grained 3D cache for high performance in a manycore context," in *Proc. IFIP/IEEE 21st Int. Conf. Very Large Scale Integr. (VLSI-SoC)*, Oct. 2013, pp. 302–307.
- [30] I. Miro-Panades, E. Beigne, O. Billoint, and Y. Thonnart, "In-situ Fmax/Vmin tracking for energy efficiency and reliability optimization," in *Proc. IEEE 23rd Int. Symp. On-Line Test. Robust Syst. Des. (IOLTS)*, Jul. 2017, pp. 96–99.
- [31] P. Meinerzhagen *et al.*, "An energy-efficient graphics processor featuring fine-grain DVFS with integrated voltage regulators, execution-unit turbo, and retentive sleep in 14nm tri-gate CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2018, pp. 38–40.
- [32] I. Bukreyev *et al.*, "Four monolithically integrated switched-capacitor DC-DC converters with dynamic capacitance sharing in 65-nm CMOS," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 6, pp. 2035–2047, Nov. 2017.
- [33] H. Meyvaert, T. Van Breussegem, and M. Steyaert, "A 1.65W fully integrated 90nm bulk CMOS intrinsic charge recycling capacitive DC-DC converter: Design; Techniques for high power density," in *Proc. IEEE Energy Convers. Congr. Expo.*, Sep. 2011, pp. 3234–3241.
- [34] T. M. Andersen *et al.*, "A 10 w on-chip switched capacitor voltage regulator with feedforward regulation capability for granular microprocessor power delivery," *IEEE Trans. Power Electron.*, vol. 32, no. 1, pp. 378–393, Jan. 2017.
- [35] T. Souvignet, B. Allard, and S. Trochut, "A fully integrated switched-capacitor regulator with frequency modulation control in 28-nm FDSOI," *IEEE Trans. Power Electron.*, vol. 31, no. 7, pp. 4984–4994, Jul. 2016.
- [36] P. Vivet *et al.*, "A 4 × 4 × 2 homogeneous scalable 3D Network-on-Chip circuit with 326 MFlit/s 0.66 pJ/b robust and fault tolerant asynchronous 3D links," *IEEE J. Solid-State Circuits*, vol. 52, no. 1, pp. 33–49, Jan. 2017, doi: [10.1109/JSSC.2016.2611497](https://doi.org/10.1109/JSSC.2016.2611497).
- [37] Y. Thonnart, P. Vivet, S. Agarwal, and R. Chauhan, "Latency improvement of an industrial SoC system interconnect using an asynchronous NoC backbone," in *Proc. 25th IEEE Int. Symp. Asynchronous Circuits Syst. (ASYNC)*, May 2019, pp. 46–47.
- [38] J. Pontes, P. Vivet, and Y. Thonnart, "Two-phase protocol converters for 3D asynchronous 1-of-n data links," in *Proc. 20th Asia South Pacific Des. Autom. Conf.*, Jan. 2015, pp. 154–159.
- [39] S. M. Nowick and M. Singh, "Asynchronous design—Part 1: Overview and recent advances," *sIEEE Des. Test. Comput.*, vol. 32, no. 3, pp. 5–18, Jun. 2015.
- [40] R. P. Reddy, A. Acharyya, and S. Khursheed, "A cost-aware framework for lifetime reliability of TSV-based 3D-IC design," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 11, pp. 2677–2681, Nov. 2020, doi: [10.1109/tcsii.2020.2970724](https://doi.org/10.1109/tcsii.2020.2970724).
- [41] C. Metzler *et al.*, "Computing detection probability of delay defects in signal line tsvs," in *Proc. 18TH IEEE Eur. TEST Symp. (ETS)*, May 2013, pp. 1–6.
- [42] C. Papameletis, B. Keller, V. Chickermane, S. Hamdioui, and E. J. Marinissen, "A DFT architecture and tool flow for 3-D SICs with test data compression, embedded cores, and multiple towers," *IEEE Des. Test. Comput.*, vol. 32, no. 4, pp. 40–48, Aug. 2015.
- [43] S. K. Goel *et al.*, "Test and debug strategy for TSMC CoWoS 2122; stacking process based heterogeneous 3D IC: A silicon case study," in *Proc. IEEE Int. Test Conf. (ITC)*, Sep. 2013, pp. 1–8, doi: [10.1109/TEST.2013.6651893](https://doi.org/10.1109/TEST.2013.6651893).
- [44] IEEE 1838 WG. Accessed: Mar. 2020. [Online]. Available: <http://grouper.ieee.org/groups/3Dtest/>

- [45] J. Durupt, P. Vivet, and J. Schloeffel, "I^JTAG supported 3D DFT using chiplet-footprints for testing multi-chips active interposer system," in *Proc. 21th IEEE Eur. Test Symp. (ETS)*, May 2016, pp. 1–8.
- [46] E. Beigne *et al.*, "A 460 MHz at 397 mV, 2.6 GHz at 1.3 V, 32 bits VLIW DSP embedding f MAX tracking," *IEEE J. Solid-State Circuits*, vol. 50, no. 1, pp. 125–136, Jan. 2015.
- [47] P.-Y. Martinez *et al.*, "ExaNoDe: Combined integration of chiplets on active interposer with bare dice in a multi-chip-module for heterogeneous and scalable high performance compute nodes," in *Proc. IEEE VLSI Conf.*, 2020.
- [48] A. Olofsson, T. Nordstrom, and Z. Ul-Abdin, "Kickstarting high-performance energy-efficient manycore architectures with epiphany," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 1719–1726.
- [49] H. Reiter, "Multi-Die IC Design Tutorial," in *Proc. 3D ASIP Conf.*, 2015, pp. 1–5.
- [50] *Calibre 3D STACK*. Accessed: 2011. [Online]. Available: https://www.mentor.com/products/ic_nanometer_design/verification-signoff/physical-verification/calibre-3dstack
- [51] zGLUE Inc. Accessed: 2014. [Online]. Available: www.zglue.com
- [52] C. Torregiani, B. Vandeveld, H. Oprins, E. Beyne, and I. D. Wolf, "Thermal analysis of hot spots in advanced 3D-stacked structures," in *Proc. 15th Int. Workshop Thermal Invest. ICs Syst.*, 2009, pp. 55–60.
- [53] T. R. Harris, P. Franzon, W. R. Davis, and L. Wang, "Thermal effects of heterogeneous interconnects on InP/GaN/Si diverse integrated circuits," in *Proc. Int. 3D Syst. Integr. Conf. (3DIC)*, Dec. 2014, pp. 1–3.
- [54] C. Santos, P. Vivet, J.-P. Colonna, P. Coudrain, and R. Reis, "Thermal performance of 3D ICs: Analysis and alternatives," in *Proc. Int. 3D Syst. Integr. Conf. (3DIC)*, Dec. 2014, pp. 1–7.
- [55] M. Graphics and W. Paper, "A complete guide to 3D chip-package thermal co-design, 10 key considerations," 2017, Tech. Rep. [Online]. Available: <https://www.mentor.com/products/mechanical/resources/overview/a-complete-guide-to-3d-chip-package-thermal-co-design-10-key-considerations-d8b0e79e-fb79-4c5a-992d-45d0d3b5f0ac>
- [56] C. Santos, P. Vivet, L. Wang, M. White, and A. Arriozaz, "Thermal exploration and sign-off analysis for advanced 3D integration," in *Proc. Design Track, DAC Conf.*, Jun. 2017.



Pascal Vivet (Member, IEEE) received the Ph.D. degree from Grenoble INPG, Grenoble, France, in 2001, designing an asynchronous microprocessor.

After four years with STMicroelectronics, Crolles, France, he joined CEA-Leti, Grenoble, in 2003, in the digital design lab. He was a Project Leader on 3D circuit design from 2011 to 2018. He is currently a Scientific Director of the Digital Systems and Integrated Circuits Division, CEA-LIST, a CEA institute. He has authored or coauthored more than 120 articles and holds several patents in the field of digital design. His research interests cover wide aspects of circuit and system level design, ranging from system integration, multicore architecture, network-on-chip, energy-efficient design, related CAD aspects, and in strong links with advanced technologies, such as 3D, nonvolatile-memories, and photonics.



Eric Guthmuller graduated from École Polytechnique and received the M.S. degree from Telecom Paris, Paris, France, in 2009, and the Ph.D. degree in computer science from University Pierre and Marie Curie (UPMC), Telecom Paris, in 2013.

He joined CEA-Leti, Grenoble, France, as a Full-Time Researcher in 2019, then with CEA-List. His main research interests include processor architectures and their memory hierarchy, in particular cache coherency for manycore and heterogeneous architectures.



Yvain Thonnart (Member, IEEE) graduated from Ecole Polytechnique, Paris, France, and received the M.S. degree from Telecom Paris, in 2005.

He joined Technological Research Division, CEA, French Alternative Energies and Atomic Energy Commission, with CEA-Leti, Grenoble, France, in 2019, then with CEA-List. He is now a Senior Expert on communication and synchronization in systems on chip, and a Scientific Advisor for the mixed-signal lab. His main research interests include networks on chip, asynchronous logic, emerging technologies integration, and interposers.



Gael Pillonnet (Senior Member, IEEE) was born in Lyon, France, in 1981. He received the master's degree in electrical engineering from CPE Lyon, Lyon, France, in 2004, and the Ph.D. and Habilitation degrees from INSA Lyon, Lyon, in 2007 and 2016, respectively.

Following an early experience as an Analog Designer in STMicroelectronics, Crolles, France, in 2008, he joined the Electrical Engineering Department, University of Lyon, Lyon. From 2011 to 2012, he held a visiting researcher position at the University of California at Berkeley, Berkeley, CA, USA. Since 2013, he has been a Full-Time Researcher at CEA-LETI, a major French research institution. His research focuses on energy transfers in electronic devices, such as power converters, audio amplifiers, energy-recovery logics, electromechanical transducers, and harvesting electrical interfaces.



César Fuguet received the M.S. degree in system's engineering from the Universidad de Los Andes (ULA), Mérida, Venezuela, in 2012, and the M.S. and Ph.D. degrees in computer science from University Pierre and Marie Curie (UPMC), Paris, France, in 2012 and 2015, respectively.

Following an experience at Kalray, Grenoble, France, he is currently a Full-Time Researcher at CEA-List, Grenoble, France. His main research interests are multicore processor architectures, cache coherency, and heterogeneous architectures with accelerators for high-performance computing.



Ivan Miro-Panades (Member, IEEE) received the M.S. degree in telecommunication engineering from the Technical University of Catalonia (UPC), Barcelona, Spain, in 2002, and the M.S. and Ph.D. degrees in computer science from University Pierre and Marie Curie (UPMC), Paris, France, in 2004 and 2008, respectively.

He worked at Philips Research, Paris, and STMicroelectronics, Grenoble, France, before joining CEA, Grenoble, in 2008, where he is currently a Research Engineer in digital integrated circuits. His main research interests are artificial intelligence (AI), Internet-of-Things, low-power architectures, energy-efficient systems, and Fmax/Vmin tracking methodologies.



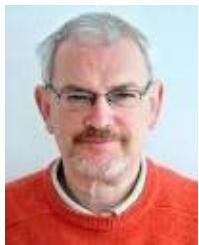
Guillaume Moritz was born in France in 1987. He graduated from Telecom Physique Strasbourg in 2010 with a specialization in micro- and nano-electronics and the associated master.

After finishing his internship at CEA-Leti, Grenoble, France, he joined Leti for two years. Then, as a subcontractor from ATOS specialized in physical design, he holds different positions with Leti, working on various advanced projects, including two major 3D circuits. He joined Leti in 2019, where he is currently focusing on physical implementation of image sensors.



Jean Durupt graduated from the École Centrale de Lyon, Lyon, France, in 1990, with a specialization in micro-electronics.

He joined CEA, Grenoble, France, in 2001, in the digital design lab. His main research interests are multicore processor architectures, circuit design, and more specifically design-for-test and circuit testability, including testability of 3D architectures.



Christian Bernard received the Engineering degree from the Grenoble Polytechnical Institute, Grenoble, France, in 1979.

After four years with Thomson, Paris, France, he worked at Bull, Paris, on mainframe HW design of CPU cores, multiprocessing, and cache coherency aspects. He joined CEA-Leti, Grenoble, in 2001, in the digital design lab. He contributed in the design of large systems of the lab covering various domains: 4G mobile baseband, space mission dedicated hardware accelerators, and many-core architectures, including the integration of cache coherency in 3D many cores. He is now retired.



Didier Varreau was born in Dôle, France, in 1954. He received the Electronic Higher Technical Diploma degree from Grenoble University, France, in 1975.

In 1976, he joined CEA-LETI, Grenoble, France, to develop instrumental electronic boards for medical and nuclear purpose. From 2003 to 2006, he worked on the FAUST project developing integrated synchronous IPs. Since 2006, he has been in charge of physical implementation of low-power energy-efficient accelerators, then since 2010, he has

been working on large multiprocessor system-on-chip, including large 3D systems. He is now retired.



Julian Pontes (Member, IEEE) is graduated in computer engineering at UEPG, Ponta Grossa, Brazil, in 2006. He received the M.Sc. and Ph.D. degrees in computer science from PUC-RS, Ponta Alegre, Brazil, in collaboration with CEA-Leti, Grenoble, France, in 2008 and 2012, respectively. His Ph.D. research work was focused on fault tolerance in asynchronous circuits and this work was extended as a PostDoc in CEA/Leti, with research contributions on 3D architecture and circuit design.

He worked with System Integration at Arm Ltd., Sheffield, U.K. He currently works with CPU design at ARM, Sophia Antipolis, France.



Sébastien Thuriès received the master's degree from the University of Montpellier, Montpellier, France, in 2003.

He joined CEA/Leti, Grenoble, France, in 2004, as a Research Engineer. He is leading the High-Density 3D Architecture and Design Group, CEA-LETI, including fine pitch 3D stacking as well as monolithic 3D (M3D). He has worked on and led several digital ASIC developments for a set of application, such as 4G digital baseband, complex imagers, system on chip, and mixed signal RF over the last decade. He has been a pioneer in FDSOI digital design and back biasing capability. He leads the research team on new architecture and design paradigm raised by M3D-IC in order to optimize the full system to technology fields.



David Coriat received the M.Sc. degree from the University of Science of Montpellier, France, in 2012.

He subsequently joined CEA-Leti Minatec, Grenoble, France. He has worked on dynamic management of power and variability in MP-SoC architectures as well as power estimation techniques in large MP-SoC architectures. His research interests now lie in low-power architectures and design.



Michel Harrand (Member, IEEE) started his career in Matra Espace, Paris, France, in 1980, where he designed automatic pilot systems for satellites. In 1985, he joined Thomson Semiconductors, Grenoble, France, where he designed numerous integrated circuits in the microprocessor, telecommunication, and mostly image compression fields, and lead a design team before being appointed as the Director of the Embedded DRAM Department in 1996. He joined CEA, Grenoble, in 2006, to prepare the creation of Kalray, Grenoble, a startup designing manycore processors, which he co-founded in 2008 as the CTO. He joined back CEA in end 2012 to explore the architecture, design, and applications of new technologies as 2.5D integrated circuits, emerging non-volatile memories, and currently neural networks. He has served in the ISSCC TPC from 2001 to 2006, and holds more than 40 patents.



Denis Dutoit joined CEA, Grenoble, France, in 2009, after working for STMicroelectronics, Crolles, France, and STEricsson, Grenoble. In CEA, he has been involved in system-on-a-chip architecture for computing and 3D integrated circuit projects. After defining the CEA-Leti's roadmap of technologies and solutions for advanced computing, he is now involved in European Projects in High Performance Computing as a Coordinator, Project Leader, and SoC Architect.



Didier Lattard was born in Saint Marcellin, France, in 1963. He received the Ph.D. degree in microelectronics from the National Polytechnic Institute of Grenoble, Grenoble, France, in 1989.

In 1990, he joined CEA-Leti, Grenoble. He was involved in the design of image and baseband processing circuits as a Senior Research and Development Engineer and Project Leader. From 2003 to 2014, he led projects in the field of NoC-based telecom and high-performance computing applications.

In 2014, he moved to the Technology Department of CEA-Tech and was involved in 3D integration projects. Since 2020, he has been leading a team developing mixed-signal circuits and software tools for near memory computing, cybersecurity, IoT, and telecom applications. He has published 60 articles in books, refereed journals, and conferences. He holds 24 patents in the fields of baseband processing, NoC architectures, and 3D integration.



Lucile Arnaud joined CEA-LETI, Grenoble, France, in 1984. She first covered design and characterization of magnetic and electromagnetic passive devices. From 2007 to 2014, she was assigned at STMicroelectronics, Crolles, France, for interconnect expertise of most advanced CMOS technology. Since 2014, she has been involved in 3D-IC developments in LETI for technology expertise and projects managing. In the last four years, she managed internal and collaborative projects for 3D interconnects development with Cu-SiO₂ hybrid bonding technologies. She authored or coauthored more than 90 articles, including some invited talks and tutorials in the IEEE conferences.



Jean Charbonnier (Member, IEEE) is graduated from the National School of Physics of Grenoble, Grenoble, France, in 2001 and received the Ph.D. degree in crystallography from the University Joseph Fourier, Grenoble, in 2006.

He joined the 3D Wafer Level Packaging Group, CEA-Leti, Grenoble, in 2008. He has been working for more than ten years in through silicon vias, 3D interconnections, and silicon interposers technology. His research interests include high-performance computing, silicon photonics interposer, as well as cryopackaging for quantum architecture applications. He is currently in charge of coordinating the High-Density 3D Integration Group, 3D Packaging Laboratory, CEA-Leti.



Perceval Coudrain received the M.S. degree in materials sciences from the University of Nantes, Nantes, France, in 2001, and the Ph.D. degree from the Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, France, in 2009.

He joined STMicroelectronics, Crolles, France, in 2002, and entered the advanced research and development group in 2005, where he was involved in the development of backside illumination and monolithic 3D integration for CMOS image sensors. For ten years, he has been focusing on 3D

integration technologies, including TSV and C-Cu hybrid bonding, and thermal management. He moved to CEA-Leti, Grenoble, France, in 2018, where his research focuses on 3D integration, fan-out wafer level packaging, and embedded microfluidics.



Arnaud Garnier graduated from INSA de Lyon, Lyon, France, in 2004. He received the Ph.D. degree from Université St Quentin en Yvelines, St Quentin, France, in 2007, in materials science in 2007 after studying the Smart Cut technology on GaN for three years in SOITEC.

He then joined CEA-LETI, Grenoble, France, to work on wafer level packaging, with a specific focus on wafer bonding, chip assembly, underfilling, 3D process integration, and advanced packaging. He currently works as a Project Leader mainly on fan-out wafer level packaging technologies.



Frédéric Berger born in Grenoble, France, in 1973. He received the B.T.S. degree from lycée Argouges, Grenoble, in 1993, in photonic optical engineering.

He started his career as a Technician in the maintenance of alarm systems, then fiber optic welders for telecommunications at Siemens/Corning. More attracted by research and development, he continues his activity in the Photonics team to develop and perfect optical amplifiers. In 2003, he joined CEA, Grenoble, as a Technician in the Packaging and Assembly Laboratory. In 2005, he participated with

SET in the development of the first FC300 equipment for 3D assemblies based on microtubes for infrared imagers. He used this technical background to carry out the assemblies of the six chiplets of the INTACT project.



Alain Gueugnot received the B.T.S. degree in microtechnology from Lycée Jules Richard, Paris, France, in 1989.

He joined CEA-DAM, Grenoble, in 1992 and then CEA-LETI at the DOPT in 2003 to work in the joint laboratory with SOFRADIR (Lynred) in the packaging. Then, he set up means of morphological characterization and metallographic expertise of assemblies of components for infrared, lighting, imager and screen using profilometers, and ionic and mechanical cross section for SEM imaging.



Alain Greiner is currently a Professor at Université Pierre et Marie Curie (UPMC), Paris, France, and an Associate Professor at Ecole Polytechnique, Paris. He was the Head of the Hardware Architecture Department, LIP6 Laboratory, from 1990 to 2010. He was the Team Leader of the public domain VLSI/CAD system ALLIANCE, and the Technical Coordinator of the SoCLib virtual prototyping platform, supported by the French Agence Nationale pour la Recherche, and jointly developed by six industrial companies and ten academic laboratories.

He is the Chief Architect of the scalable, shared memory, and manycore TSAR architecture, and is presently working on scalable operating systems for those kinds of machines.



Quentin L. Meunier received the Diploma degree from the Ensimag School, Grenoble, France, in 2007, and the Ph.D. degree in computer science from the Université de Grenoble, Grenoble, in 2010.

Since 2011, he has been an Associate Professor at the LIP6 Laboratory, Sorbonne Université, Paris, France. His research interests include many core architectures and cache coherence, high-performance computing, and side-channel attacks and counter-measures.



Alexis Farcy graduated in electronic engineering from the "Institut des Sciences et Techniques de Grenoble," France, in 2000, and the Ph.D. degree in electronic, optronic and systems from the University of Savoie, Chambery, France, in 2009.

He was employed by STMicroelectronics, Crolles, France. From 2000 to 2007, he was among the Advanced Interconnects and Passive Components Module, focusing on interconnect performance analysis for advanced technology nodes, integration of advanced inductors and 3D capacitors in BEOL, and high-frequency characterizations of low- k and high- k dielectrics. Since 2007, he has been in the field of 3D integration on innovative technologies, processes and materials for 3D integration, and performance assessment for photonics and image sensors.



Alexandre Arriordaz received the master's degree in electronics from the University de Nice-Sophia-Antipolis, France.

He is a Senior Product Engineering Manager for caliber design solutions at Mentor – A Siemens Business, Montbonnot, France. He is leading product management and software development teams located, Grenoble, France, focusing on circuit reliability and verification product lines. In parallel to this activity, he is also a technical interface for various European projects dealing with research and development topics, such as 3D-IC or silicon photonics. Prior to joining Mentor, he was a Full-Custom Design Engineer at Freescale Semiconductor (now NXP), Grenoble, working on advanced testchip/SRAM compiler developments.



Séverine Chéramy (Member, IEEE) received the Engineering degree from Polytech Orleans, Orleans, France, in 1998, having specialized in material science.

She has spent over eight years at GEMALTO, Aix en Provence, France, a leading smart-card company developing technologies for secure solutions, such as contactless smart cards and electronic passports. In 2008, she joined CEA-Leti, Grenoble, France, as a 3D Project Leader and then as a 3D Integration Laboratory Manager. This group develops technology and integration for 3DIC, in strong relationship with 3D design, model, and simulation teams. Since January 2017, she has been responsible for 3DIC integration strategy and related business development. She is also the Director of the 3D project of the Institute of Technological Research (ITR) Nanoelec.



Fabien Clermidy (Member, IEEE) received the Ph.D. and Thesis Supervisor degrees from INPG, Grenoble, France, in 1999 and 2010, respectively.

In 2000, he joined CEA-LIST, Paris, France, where he was involved in the design of an application-specific parallel computer. In 2003, he joined CEA-LETI, Grenoble, in the digital circuit laboratory, where he led the design of various large many-core circuits. He is currently the Head of the Digital System and Circuit Division, CEA-LIST, a CEA institute. He has published more than 80 articles in international conferences. He holds 14 patents. His research interest covers wide scope of digital systems: many-core architecture, network-on-chip, energy-efficient design, embedded systems, and interaction with advanced technologies.