

Architecture, Chip, and Package Codesign Flow for Interposer-Based 2.5-D Chiplet Integration Enabling Heterogeneous IP Reuse

Jinwoo Kim^{1b}, Graduate Student Member, IEEE, Gauthaman Murali^{1b}, Heechun Park, Eric Qin, Graduate Student Member, IEEE, Hyoukjun Kwon^{1b}, Graduate Student Member, IEEE, Venkata Chaitanya Krishna Chekuri^{1b}, Graduate Student Member, IEEE, Nael Mizanur Rahman, Graduate Student Member, IEEE, Nihar Dasari, Arvind Singh, Member, IEEE, Minah Lee^{1b}, Graduate Student Member, IEEE, Hakki Mert Torun^{1b}, Graduate Student Member, IEEE, Kallol Roy, Member, IEEE, Madhavan Swaminathan, Fellow, IEEE, Saibal Mukhopadhyay, Fellow, IEEE, Tushar Krishna, Member, IEEE, and Sung Kyu Lim, Senior Member, IEEE

Abstract—A new trend in system-on-chip (SoC) design is chiplet-based IP reuse using 2.5-D integration. Complete electronic systems can be created through the integration of chiplets on an interposer, rather than through a monolithic flow. This approach expands access to a large catalog of off-the-shelf intellectual properties (IPs), allows reuse of them, and enables heterogeneous integration of blocks in different technologies. In this article, we present a highly integrated design flow that encompasses architecture, circuit, and package to build and simulate heterogeneous 2.5-D designs. Our target design is 64-core architecture based on Reduced Instruction Set Computer (RISC)-V processor. We first chipletize each IP by adding logical protocol translators and physical interface modules. We convert a given register transfer level (RTL) for 64-core processor into chiplets, which are enhanced with our centralized network-on-chip. Next, we use our tool to obtain physical layouts, which is subsequently used to synthesize chip-to-chip I/O drivers and these chiplets are placed/routed on a silicon interposer. Our package models are used to calculate power, performance, and area (PPA) and reliability of 2.5-D design. Our design space exploration (DSE) study shows that 2.5-D integration incurs 1.29× power and 2.19× area overheads compared with 2-D counterpart. Moreover, we perform DSE studies for power delivery scheme and interposer technology to investigate the tradeoffs in 2.5-D integrated chip (IC) designs.

Index Terms—2.5-D integrated chip (IC), chiplet, electronic design automation (EDA) flow, interposer, power, performance, and area (PPA), reliability.

Manuscript received February 16, 2020; revised July 2, 2020; accepted July 18, 2020. Date of publication August 24, 2020; date of current version October 23, 2020. This work was supported by the Defense Advanced Research Projects Agency (DARPA) Common Heterogeneous Integration and IP Reuse (CHIPS) Program under Award N00014-17-1-2950. (Corresponding author: Jinwoo Kim.)

Jinwoo Kim, Gauthaman Murali, Heechun Park, Eric Qin, Hyoukjun Kwon, Venkata Chaitanya Krishna Chekuri, Nael Mizanur Rahman, Nihar Dasari, Minah Lee, Hakki Mert Torun, Madhavan Swaminathan, Saibal Mukhopadhyay, Tushar Krishna, and Sung Kyu Lim are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: jinwookim@gatech.edu; limsk@ece.gatech.edu).

Arvind Singh is with the Rambus Cryptography Research Division, Rambus Inc., San Francisco, CA 94105 USA.

Kallol Roy is with the Institute of Computer Science, University of Tartu, 50090 Tartu, Estonia.

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVLSI.2020.3015494

1063-8210 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See <https://www.ieee.org/publications/rights/index.html> for more information.

I. INTRODUCTION

AS THE process technology continuously scales and the design complexity increases, traditional 2-D-based integrated chip (IC) design may no longer catch up with the scaling trend of Moore's law [1]. Moreover, 2-D IC design becomes difficult to satisfy the fast growing demands of high performance from the market. Many researchers found solutions from the concept of 3-D IC design [2]. It redeemed the shortcomings of 2-D IC design by stacking the chips vertically and connecting them with shorter interconnect and high bandwidth. However, 3-D IC design exposed some weaknesses, such as large overhead from through-silicon via (TSV) and high temperature issues [3].

Following that, interposer-based 2.5-D IC design is proposed to overcome the problems, as well as maintain the strengths, of 3-D IC [4]. Instead of stacking chips vertically, all chips are placed on the interposer side-by-side, and connected through the interposer with high speed and bandwidth. It eliminates the use of TSVs in the chip, and avoids the thermal problem caused by high vertical power density of 3-D ICs. Moreover, FOVEROS technology from Intel and Zen 2 microarchitecture from AMD indicates that 2.5-D IC technology is no longer an alternative but a new trend in system-on-chip (SoC) design.

Interposer-based 2.5-D IC design allows block-level heterogeneous integration, which means that all functional circuit blocks are designed separately under different environments and integrated, rather than designed and fabricated monolithically into a single SoC. Fig. 1 shows a conceptual view of an interposer-based 2.5-D IC and its cross section view. The 2.5-D IC has an interposer on top of the package and the functional blocks, named *chiplets*,¹ are mounted on the interposer. All connections between chiplets are made through the interposer to achieve high speed and throughput.

2.5-D chiplet integration provides promising features of heterogeneity, reusability, and easy update of intellectual prop-

¹A chiplet is defined as a functional module that contains interposer I/O drivers and in its bare die form with microbumps on the bottom to be mounted on an interposer to communicate with other chiplets.

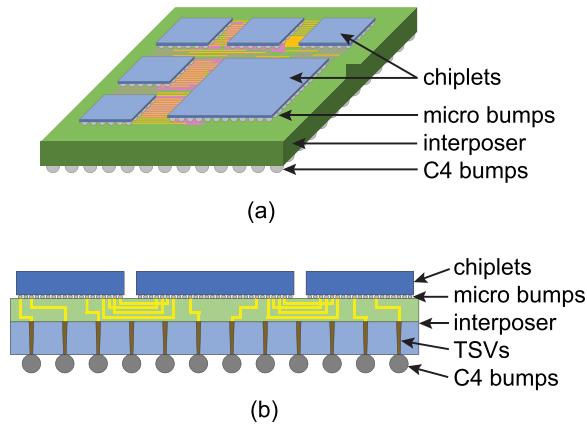


Fig. 1. 2.5-D chiplet integration with an interposer. (a) Interposer-based 2.5-D IC. (b) Cross-sectional view of 2.5-D IC.

erties (IPs) in SoC design, compared to a traditional 2-D IC design approach. With this architecture, each IP can be independently designed into a chiplet under its most suitable technology node and assembled into the SoC. This design approach enables SoC designers to simply choose appropriate off-the-shelf chiplets and heterogeneously integrate them into the target SoC, which drastically reduces the design time, complexity, and cost by reutilizing predesigned chiplets as plug-and-play modules. Moreover, the development risk of SoC in 2.5-D integration becomes significantly lower than a traditional 2-D IC design because the known good dies (KGDs) are selected as chiplets [5]. Besides, the system update is greatly simplified because it only needs to swap out chiplets that are necessary, instead of redesigning the entire SoC from scratch.

In this article, we first present our Reduced Instruction Set Computer (RISC)-V-based 64-core architecture named ROCKET-64 [6] for chiplet integration. Next, we present a vertically integrated electronic design automation (EDA) flow for chiplet creation and integration, which covers the design phases of architecture, circuit, and package. We also present a new logical protocol called hybrid-link to reduce overheads of 2.5-D IC design. Moreover, we provide power, performance, and area (PPA), signal integrity (SI) and power integrity (PI) data of 2.5-D IC design for design space exploration (DSE) with quantitative comparisons. We choose a target design of ROCKET-64 with network-on-chip (NoC) configuration to show stepwise explanation of the overall flow.

We claim the following contributions: 1) our 64-core RISC-V architecture is scalable and appropriate for chiplet integration; 2) we propose a new logical protocol that is well fitted for 2.5-D IC design; (3) we generate interposer-based 2.5-D design including interposer routing and the layout of each chiplets with optimized I/O drivers by using commercial tools; 4) we analyze power delivery network (PDN) of 2.5-D IC to show the time and frequency domain characteristics; 5) we analyze PPA of interposer-based 2.5-D ICs and compare the results with monolithic 2-D IC to investigate overheads of 2.5-D design; and 6) we analyze tradeoffs in 2.5-D IC designs depending on power delivery scheme and interposer technology. To our best knowledge, this is the first work to fully quantify the design gap, which enables DSE of various

aspects in terms of PPA, SI, and PI using GDS layouts and sign-off simulations.

II. RELATED WORK

Before applying 2.5-D technology to real designs, thorough analysis on 2.5-D IC design should be preceded by various perspectives. There are existing studies on 2.5-D IC design focused on the design methodology or utility point of view such as analysis of design cost aspect [7] and bump assignment algorithm for 2.5-D interposer design [8]. However, there is no numerical analysis in terms of the actual PPA of 2.5-D design in these works.

Recently, some researchers have explored the codesign methodology for 2.5-D IC covering chip to package including [9]. Kabir and Peng [9] have proposed 2.5-D design flow which design 2.5-D package together with chiplets in the same design environment. They first synthesize the gate-level netlist of the entire system and perform the architecture-aware partitioning to subdivide the system into multiple chiplets. Chiplet designs are done in a single design environment using a hierarchical design scheme, the package routing by RDL planner, all chiplet designs and package routing are assembled in one single design at the end. This work has also claimed that its analysis results are more accurate and reliable because it reflects the parasitics of the entire chip-package system.

However, despite the presence of key features, this study presents some limitations. As their flow starts with the synthesis of target design, a huge resource is required for a very large architecture such as, ROCKET-64. Besides, this flow is missing the heterogeneity of 2.5-D integration because all flows are performed in a single design environment. It is also overlooked that the inductance should be considered in SI because the dimensions of package wires are larger than on-chip.

As the organic interposer has introduced as an emerging technology to replace the silicon interposer due to its benefit of fabrication cost, researchers have investigated the characteristics of these interposer technologies and their tradeoffs. However, these studies have focused on SI and PI from a given substrate technology [10], [11]. They have neither carried out their studies at full-system level nor provide detailed PPA comparisons with other substrate technologies. Moreover, the tradeoffs between silicon and organic interposers are generally well known, however, a thorough and quantitative analysis in the system level has not been performed.

In this article, we focus on the EDA flow of 2.5-D IC design which covers the entire system level. This work significantly extends the prior work [6] by expanding the analysis area to SI and PI of the interposer. Moreover, we provide the quantitative analysis results with commercial-grade layouts which enables a realistic DSE of 2.5-D IC design. We demonstrate that our flow is applicable to various technologies by analyzing tradeoffs according to the power delivery configuration and interposer technology at the end of this work.

III. ARCHITECTURE AND DESIGN SETTING

A. RISC-V 64-Core Architecture

We create a 64-core architecture named ROCKET-64 based on RISC-V Rocketcore [12] as our benchmark design targeting

graph algorithm and computing. In our design, we have divided the entire SoC into chiplets considering the reusability of each IP and easy updates of the system, which are the key features in 2.5-D integration. We have designed the core module and L2 cache memory module of RISC-V as separate chiplets, therefore, the upsizing of memory capacity and updating the core architecture are available in a simple way as plug-and-play. Moreover, we also have generated an NoC and a memory controller (MC) as chiplets for the same reason.

As shown in Fig. 2, our ROCKET-64 consists of eight Rocket tiles, a centralized NoC as an arbiter, a 4-channel MC to access external DRAMs, and four integrated voltage regulators (IVRs) as power management modules which convert 3.6–0.9 V and provide maximum 12 A of current to our benchmark design. Each Rocket tile consists of an octacore RocketCore and L2 cache memory module. Each module contains I/O drivers only for 2.5-D interposer design.

Our centralized NoC consists of 12 routers interconnected in a 4×3 mesh topology. Links from each Rocket tile and MC are connected to the external ports of routers. Each router has five ports (N, E, S, W, and external) with four virtual channels at each port. The router implementation is based on an one-cycle pipeline design, which consumes one cycle in the router logic and additional one cycle for link traversal, used in OpenSMART [13]. We implement matrix arbiters that provides fairness for input virtual channel arbitration and switch allocation to prevent starving at any core.

B. Overall EDA Flow

Fig. 3 shows the overall flow of our chiplet creation and integration. Our EDA flow takes interposer PDK, design netlist, logical protocol, and chiplet PDK as the initial inputs, generates the layouts of interposer and each chiplet, and performs timing, PPA, interposer PDN analysis with existing commercial tools for each step.

In an interposer design step, we generate the layout of interposer including the footprint of each chiplet and the routing information between chiplets. We extract the wirelength distribution of interposer wires for timing analysis. The interposer channel with corresponding dimensions is characterized using a full-wave EM solver and Ansys HFSS. Next, S-parameters of interposer wires defining the impedance and coupling profile are extracted. These are then converted to SPICE models using the broadband SPICE generator of keysight ADS.

Moreover, we create the interposer PDN model using transmission matrix method (TMM) [14] for frequency and time domain PDN analysis. We use a lumped Π model which consists of resistance, inductance, conductance, and capacitance (RLGC) values and perform MATLAB simulation to analyze the PDN impedance and the transient response of IVR.

Based on silicon interposer design, we design I/O drivers to handle up to 10-mm length of interconnections in the interposer layer [15]. With well-designed I/O drivers, we generate the layouts of chiplets in chiplet design step. We use Cadence Innovus to perform place-and-route (P&R) of chiplets with usual 2-D design method. We analyze PPA of interposer-based 2.5-D design in the final step using Synopsys PrimeTime.

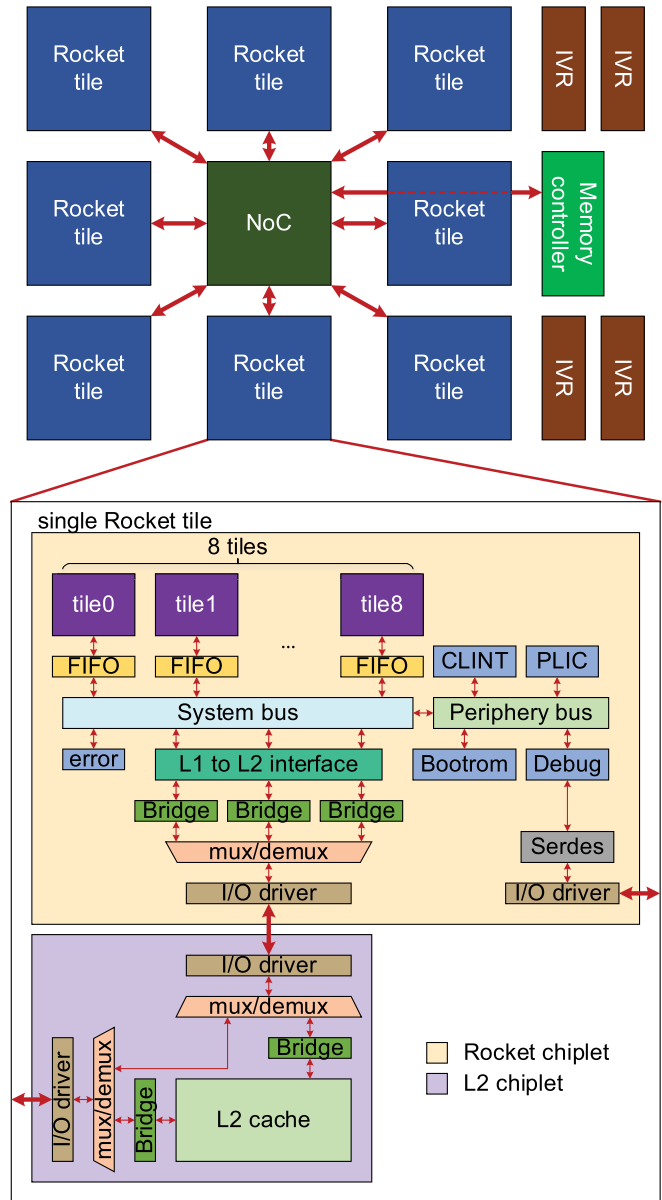


Fig. 2. Our proposed 64-core architecture for chipletization and 2.5-D integration.

Full-chip timing and power analysis for individual chiplets is straightforward after their layouts are constructed in chiplet P&R step. Once our interchiplet I/O drivers are built and chosen to handle the given interconnect length, we calculate their delay and power consumption using their SPICE models and interposer wire models. We then add these values to chiplet delay and power data. Our interposer interconnects are pipelined due to the flip-flops used in the I/O drivers, which simplify timing calculation for the entire interposer design.

C. Interposer Design Rules

In the past few years, as the design complexity of a single module increases, dense interposer designs with fine pitch of RDLs and microbump have been required in heterogeneous integration due to high I/O counts and the increasing number

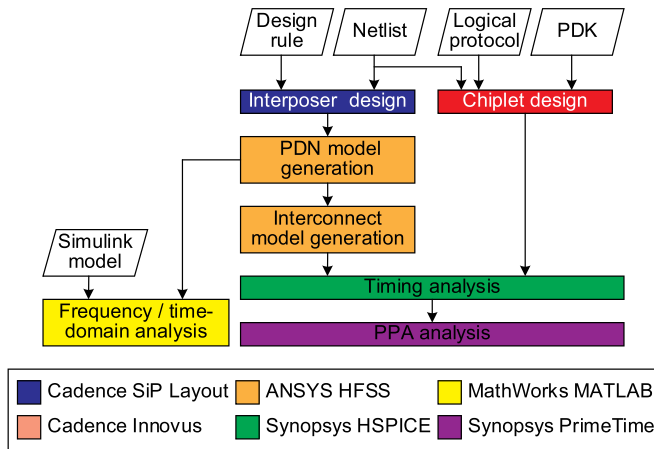


Fig. 3. Our EDA flow using commercial tools.

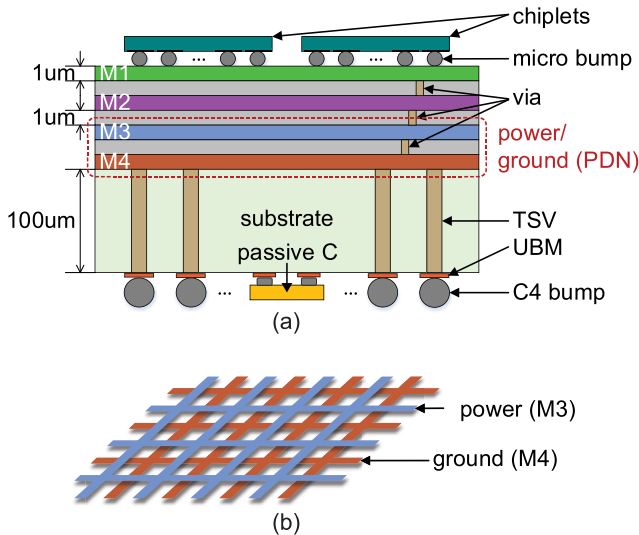


Fig. 4. Vertical stack-up of our interposer-based 2.5-D IC. (a) Vertical stack-up. (b) Mesh-type PDN.

of interconnections between chiplets. A representative example of satisfying these requirements is a silicon interposer. Taiwan Semiconductor Manufacturing Company (TSMC), Limited and Xilinx, Inc., have suggested Chip-on-Wafer-on-Substrate (CoWoS) technology [16] which provides minimum $0.8\text{-}\mu\text{m}$ pitch RDLs and supports over 200k of microbumps with $45\text{-}\mu\text{m}$ microbump pitch. They have demonstrated Virtex-7 2000T FPGA, which consists of four different 28-nm FPGA dies and has more than 10 000 die-to-die connections, as the application of CoWoS.

The design rules for our interposer design in this article are shown in Table I and Fig. 4 based on TSMC CoWoS. We choose the silicon interposer with $0.8\text{-}\mu\text{m}$ fine pitch RDLs and $40\text{-}\mu\text{m}$ -pitch microbumps for our benchmark.

IV. CHIPLETIZATION RESULTS

For the interposer-based 2.5-D IC design, we first divide a single SoC into multiple functional blocks. We use the natural IP boundaries—core, cache, NoC, and MC to create a total of 22 chiplets and eight passive components. Moreover,

TABLE I
DESIGN RULES FOR OUR SILICON INTERPOSER BASED ON TSMC CoWoS TECHNOLOGY

Design rule	Value
Metal layer #	4
Metal thickness	$1\mu\text{m}$
Dielectric thickness	$1\mu\text{m}$
Min. line width / spacing	$0.4\mu\text{m} / 0.4\mu\text{m}$
Via size	$0.7\mu\text{m}$
Through Via size / depth	$10\mu\text{m} / 100\mu\text{m}$
Die-to-die spacing	$100\mu\text{m}$
micro-bump pitch	$40\mu\text{m}$
C4 bump pitch	$400\mu\text{m}$
PDN width / spacing	$40\mu\text{m} / 100\mu\text{m}$

we add IVR chiplets, embedded inductors, and low-profile capacitors for the efficient power delivery to chiplets on the interposer. Before generating chiplets from these functional blocks, two design features must be strongly considered: an interface protocol and I/O drivers.

A. 2.5-D Interface Protocol

1) *Interface Protocol Comparison*: The study of interface protocols for systems with modular IP blocks is important for easy system design, integration, and verification. On-chip IPs today use a rich set of protocols; examples include AMBA advanced extensible interface (AXI) or its variants such as AXI-lite and ACE used by ARM-based IPs, TileLink used by RISC-V based IPs, Avalon used by Intel/Altera, and so on. Unfortunately, these cannot be ported directly to chiplets as they have hundreds of I/O signals to support address, data, and commands for multiple individual channels. Wires are relatively cheap as on-chip since the area of an IP block is dominated by logic, not I/O, since the minimum wire pitch in modern technology nodes is $0.09\mu\text{m}$. For a chiplet, however, C4 bumps to connect to the interposer are much wider such as $180\mu\text{m}$, and can potentially completely dominate the area of a chiplet.

2) *Hybrid-Link*: In this work, we propose a new protocol called Hybrid-Link. Hybrid-Link is designed keeping three goals in mind.

- 1) A need for a standard protocol applicable across different chiplets.
- 2) 2.5-D ICs should have low number of external I/Os.
- 3) Different chiplets have different communication requirements.

Our new interface protocol is tailored to 2.5-D integrations with its low I/O overhead and lightweight/extended protocol mode. Although the concepts of Hybrid-Link are similar to AXI4, Figs. 5 and 6 show our study of the ideal flit size range and the functionality requirements of different chiplets in our benchmark design.

In the case of Rocket chiplet, the logic area overshadows the physical channel overhead. It means that the I/Os are not contributing to additional area. However, in the case of NoC chiplet, there is huge microbump area cost even with very narrow physical channel width. This is because

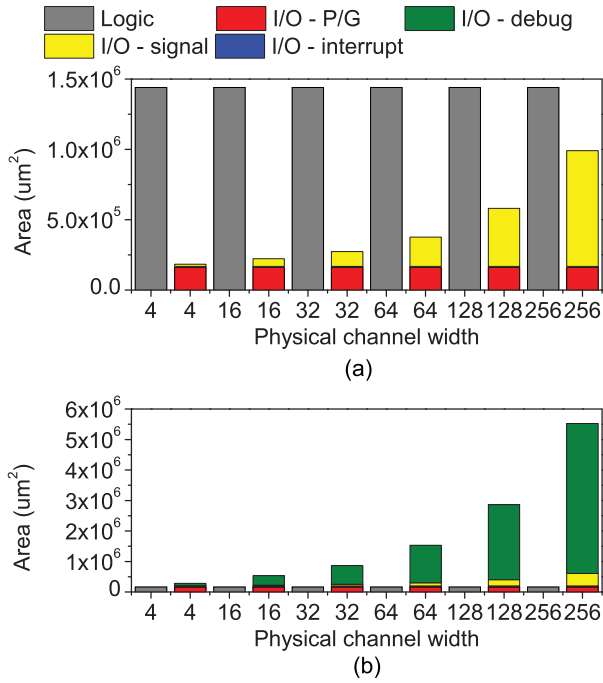


Fig. 5. Relationship between the size of chiplet versus I/O counts. (a) Rocket tile chiplet. (b) NoC chiplet.

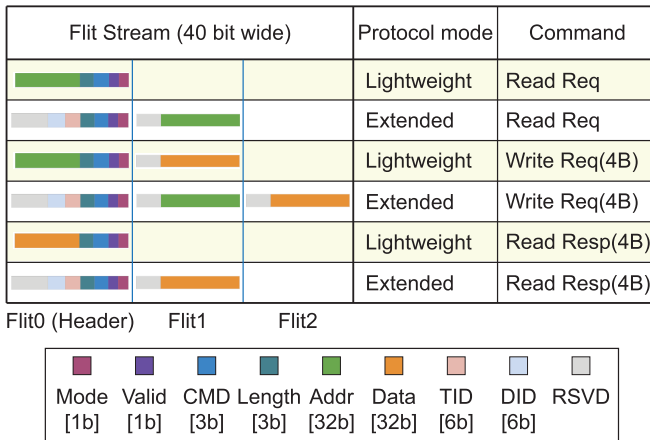


Fig. 6. Flit representation of Hybrid-Link.

NoC contains numerous Hybrid-Link I/O ports along with much smaller logic overhead than Rocket chiplet. A narrow interface protocol like Hybrid-Lite for 2.5-D ICs is necessary to keep the chiplet area reasonable, and not let I/O bump area dominate. Moreover, Hybrid-Link's 40 b interface can help design smaller chiplets (i.e., smaller logic area) without incurring an area penalty due to I/O.

Fig. 6 shows a sample flit² representation of common commands. Hybrid-Link uses a default flit width of 40 bits—though this can be further reduced, at the cost of serialization. The protocol can operate in two modes—lightweight and extended.

The lightweight mode is for simple point-to-point connections, e.g., a video filtering chiplet streaming data to an SRAM

²A flit is the number of bits of data transfer over the physical link.

chiplet. In this mode, the protocol provides a few bits for command, while the rest of the bits are used by address and data. As shown in Fig. 6, lightweight mode requires only one flit for read requests and responses, and two-flits for write requests. In the extended mode, more complex transactions (such as coherence transactions from CPU to memory via the NoC) can be supported. The extended mode provides fields for destination and transaction identifiers (DID and TID) to support AXI transactions.

The extended mode also supports multiple virtual channels to allow better buffer utilization and provide deadlock freedom. Additional communication features may be added to the RSVD field. There is one protocol bit in the header flit that determines whether the packet will be read in lightweight or extended mode. A finite-state machine will determine how to parse the following flits fields based on protocol bit. Both protocol modes allow variable packet lengths and common commands. ROCKET-64 uses the extended mode for the Rocket, L2, NoC chiplets and MC chiplets.

B. Bridges and I/O Drivers

To translate common interface protocols, such as AXI4 and TileLink to Hybrid-Link, we implemented FIFO queues and bridge FSMs. The FIFO queues are used to store common flit fields across the two prototypes, and the FSMs are used to remap the field representation to Hybrid-Link and vice versa. The FSMs are also responsible for flit arbitration and ready signals handling. The bridge consumes negligible area compared to the size of rocket chiplet.

Moreover, chiplet-to-chiplet interconnections are generated through the interposer layer which has larger pitch and longer wirelength compared to monolithic 2-D design, so additional I/O drivers are necessary for each input and output to drive the signals without any loss. In this work, we choose Intel's Advanced Interface Bus (AIB) [17] as our I/O driver model.

We design AIB based as shown in Fig. 7 on [15], because it is essential to design I/O driver optimized for wirelengths in 2.5-D IC designs to achieve high data rates. Our tool supports two optimization modes which are the delay-opt and the power-opt for the design purpose. In the delay-opt mode, our I/O driver design tool selects the driver/receiver (Tx/Rx) pair with the minimum propagation delay, and the Tx/Rx pair with the minimum power consumption in the power-opt mode at the end of the optimization process.

In the overall flow, we first choose the length from the distribution of interposer wirelength as the target wirelength and implement it as SPICE subcircuit of a parameterized transmission line model. A SPICE netlist is then generated for entire system similar to [15]. Finally, the netlist is simulated in HSPICE over wide search space of Tx/Rx sizes and the combination resulting in the minimum propagation delay or power consumption is chosen depending on the selected optimization mode. The Verilog netlists for the resulting Tx/Rx sizes are then generated using a register transfer level (RTL) template for the I/O macros.

Table II shows the AIB optimization results in both delay-opt and power-opt modes. For the average wirelength of the

TABLE II

AIB OPTIMIZATION RESULTS OF TWO OPTIMIZATION MODES FOR THE AVERAGE WIRELENGTH IN THE INTERPOSER DESIGN

Optimization mode	Delay-opt	Power-opt
Target wirelength (μm)	3,568.9	
Driver (Tx) size	$\times 80$	$\times 16$
Receiver (Rx) size	$\times 16$	$\times 4$
Propagation delay (ps)	60.2	135.4
Power consumption (μW)	171.6	157.7

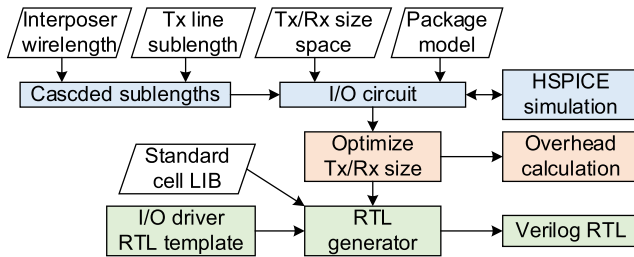


Fig. 7. AIB design and optimization flow.

interposer design in Fig. 14(a), the size of Tx/Rx pair has been chosen as 80/16 which results 60.2 ps of the propagation delay and 171.6 μW of the power consumption in the delay-opt mode. In the power-opt mode, 5 \times smaller Tx and 4 \times smaller Rx are chosen when compared to the delay-opt mode. As the smaller Tx/Rx pair is selected in the power-opt mode, the propagation delay is 2.2 \times longer than the delay-opt mode, however, it shows 8.1% saving in power consumption. As the SI should be guaranteed at least, only 8.1% power saving is achieved in the power-opt mode. In this article, our main priority is to ensure SI at high/core frequencies. Therefore, we have chosen the delay-optimized designs in this article as the resulting power/energy overhead is also within the design budget.

C. IVR and Embedded Inductor

We use the IVR chiplet presented in [18] which design consists of a power stage, feedback/control loop and an LC output filter as shown in Fig. 8. The output filter of the power stage is implemented using an inductor and a capacitance. The feedback loop of IVR consists of ADC, type-III proportional integrate-differential (PID) controller and a digital pulsewidth modulation (DPWM) block. Based on the voltage error from a reference, the compensator output is fed to a DPWM engine, generating gate signals with a duty cycle based on control word. The dc-dc conversion is achieved by duty-cycling the “ON-OFF” period of the power stages, and regulation is performed by changing the duty-cycle.

We choose the solenoidal inductor with Nickel-Zinc (NiZn) ferrite magnetic core [19] and integrate on the top metal layer of silicon interposer in our 2.5-D design. The co-optimization with IVR design [20] is formulated as a multiobjective problem to determine inductor geometry, switching frequency, and output capacitance. The goal is to find such design parameters to find the optimal trade-off between power conversion

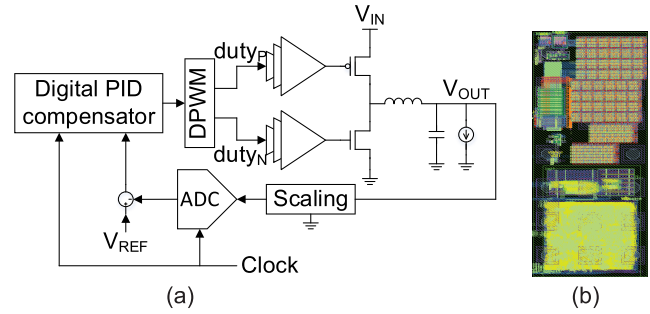


Fig. 8. (a) Block diagram and (b) GDS layout of our IVR chiplet implemented in a commercial 130-nm technology.

efficiency, voltage droop, voltage droop, settling time, and the area of inductor. Our embedded inductor is designed to be 25 nH with 3 A of the saturation current due to the limited area of the silicon interposer.

D. Chiplet Layouts

We perform chiplet P&R using Cadence Innovus as the physical design tool with selected protocol translator and AIB. We first run the microbump assignment. As chiplets are mounted on an interposer and connected to each other through microbumps, the bump assignment is an important factor to optimize the length of signal interconnection. We choose the regular bump assignment which places power and ground (P/G) microbumps at the periphery and signal bumps in the center of the chiplet as shown in Fig. 9.

Each chiplet has a minimum 100 P/G microbumps on the chiplet to ensure a worst case PDN dc resistance of below 15 m Ω . However, in the case of a chiplet with an aspect ratio more than 2 such as NoC and MC chiplets, we move or insert additional P/G microbumps at the middle of the chiplet to avoid a high IR-drop in chiplet power rail. Moreover, as Rocket chiplet consumes the most power among the chiplets, we add additional P/G bumps for sufficient power delivery.

With well-defined microbump assignment, we use area I/O placement [21] in our chiplets to minimize the distance from AIB to the signal bump. The tool handles AIBs as macrocells and places them on the proper positions to meet the timing constraint.

The chiplet list of our benchmark design and their GDS layouts with 1-GHz target frequency are shown in Table III and Fig. 10. We use a commercial 28 nm which supply level of 0.9 V for all chiplets except IVR chiplets with a commercial 130 nm which supply level of 1.2 V as technology nodes for our chipletization.

V. INTERPOSER-BASED 2.5-D IC DESIGN

A. Interposer Design Results

The process of designing the interposer consists of C4 bump assignment according to the interposer floorplan and placement of chiplet dies and interposer routing. As shown in Table III and Fig. 11, the most of external connections come from MC chiplet which is placed in the center of interposer.

TABLE III
CHIPLET LIST IN OUR ROCKET-64 BENCHMARK DESIGN

Chiplet	I/O bump #			Signal bump #			Footprint (mm × mm)	Bump array	Technology node
	Total	Signal	P/G	Internal	External	Common			
Rocket	441	58	383	46	10	2	1.70 × 1.70	21 × 21	commercial 28nm
L2 cache	196	92	104	90	-	2	1.46 × 1.46	14 × 14	commercial 28nm
NoC	663	554	109	552	-	2	0.68 × 1.56	17 × 39	commercial 28nm
Memory controller	700	587	113	185	400	2	0.80 × 1.40	20 × 35	commercial 28nm
IVR	360	9	351	-	9	-	0.48 × 1.20	12 × 21	commercial 130nm
Passive L	-	-	-	-	-	-	1.60 × 3.40	-	embedded inductor
Passive C	-	-	-	-	-	-	1.20 × 0.70	-	silicon capacitor

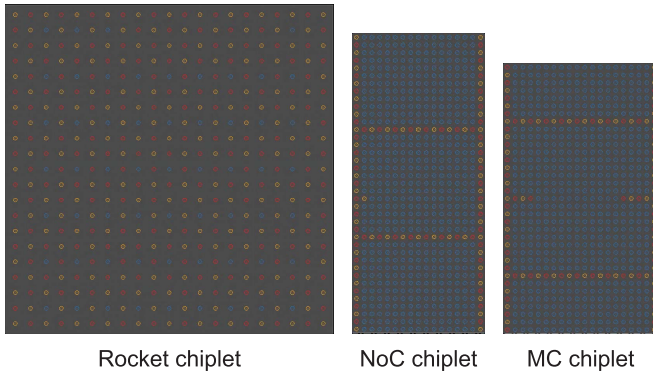


Fig. 9. Regular bump assignment in chiplet designs. Blue: signal bumps. Red: power bump. Yellow: ground bump.

C4 bump assignment is also important to reduce the length of external connection, therefore, we place C4 signal bump of the interposer to be as identical as microbump assignment of MC chiplet. With C4 and microbump assignments, we generate die data of both chiplets and the interposer, which contain bump coordinate and type, from Verilog netlists as input for floorplanning and interposer routing.

GUI-based floorplanning and interposer routing have been done by using Cadence SiP Layout XL. We first set up technology file including metal stack and via structures which provides physical and electrical information based on Table I. By importing die data into the tool, we place all the dies of chiplets on the interposer for the next routing step. In our benchmark design, we placed passive capacitors at the bottom of the interposer to reduce entire footprint as shown in Fig. 11(b). Automatic Router provided by the tool is used for 1420 interconnections in interposer layer and performs Manhattan routing same as on-chip routing as shown in Fig. 12. As we use Manhattan routing, M1 and M3 layers are used for the vertical routing and M2 and M4 layers for the horizontal routing.

While in the routing step, the data skew problem should be considered an important factor. Unlike monolithic 2-D ICs, the wire length of the signal between chiplets in 2.5-D system can reach several millimeters in the case of nonneighboring connections. Due to the distance differences between bump pairs in the single bus, each signal can arrive at its destination with different timing. Especially in the case of nonneighboring connection where source and sink chiplets are placed far apart, this problem should be highly critical in interposer

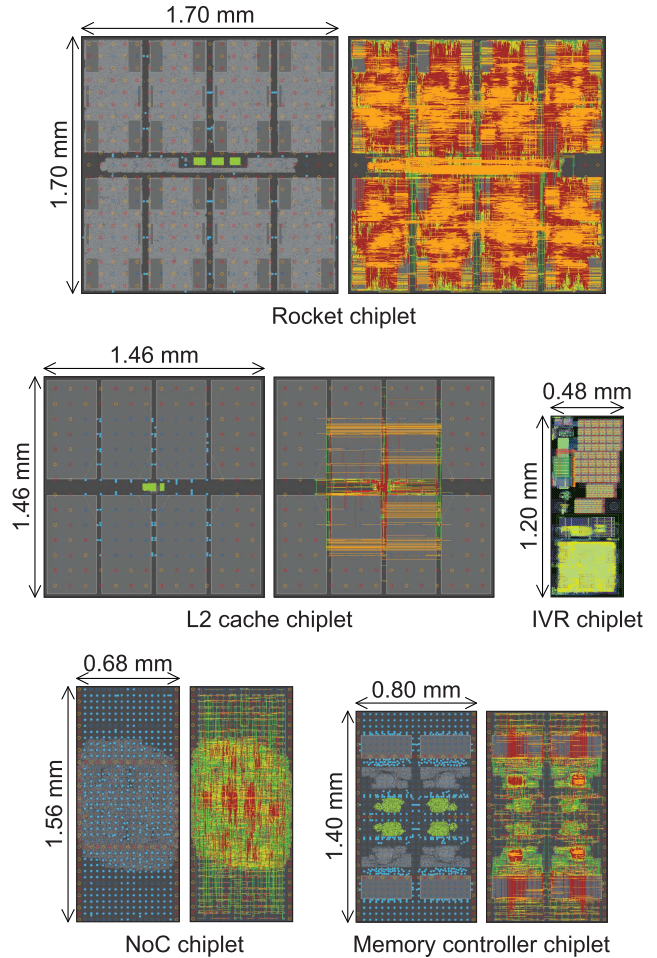


Fig. 10. Commercial 28- and 130-nm physical layouts of the chiplets in our ROCKET-64 architecture. Green part: protocol translator/bridge logic. Blue part: AIB drivers.

routing. To avoid it, we added a design constraint, named match group (MG).

The new design constraint creates a new design rule that causes wire lengths or propagation delays of signals to be in the specified target distribution for signals belonging to the same group. Compared to when MG is applied to one of our benchmark design buses and when MG is not, the wire length variation is reduced from 6960 to 500 μm as shown in Fig. 13. We assign each bus in our design as each MG with a design constraint of 500 μm , which causes a delay difference less than 5 ps.

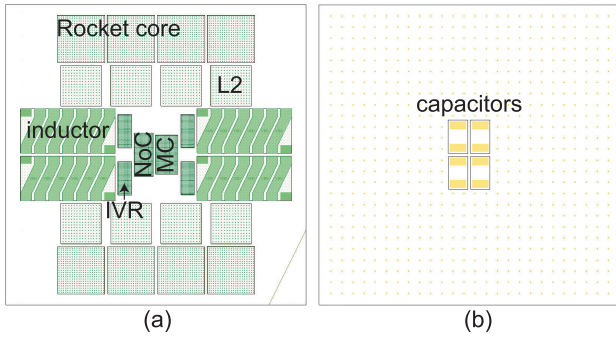


Fig. 11. Floorplan of our silicon interposer. (a) Top and (b) bottom sides.

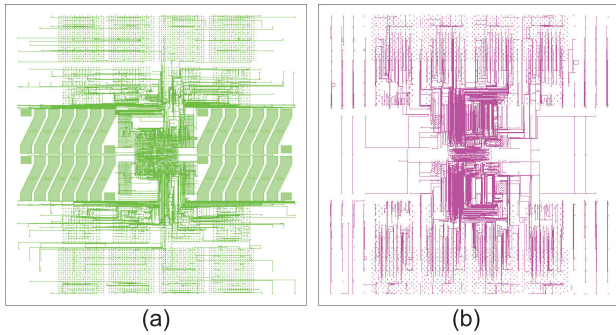


Fig. 12. (a) M1 and (b) M2 metal layers in the silicon interposer design.

Our silicon interposer design results are shown in Table IV and Fig. 14. 1420 nets are routed on the silicon interposer layer and four metal layers are used in order to demonstrate the 2.5-D design of our benchmark including PDN.

B. Interposer Timing and Power Analysis

We consider AIB with full-swing signal as our I/O drivers. A strong output driver is required to drive long interposer wires. Moreover, interposer wires have significant inductance leading to signal reflections from both driver and receiver ends due to their large dimensions. To eliminate these reflections, the impedance of final driver stage is matched to the characteristics of the package wire. To reduce overheads of the I/Os, I/O driver runs at full-swing of the supply voltage. For a commercial 28-nm technology node, the final driver size is chosen to be $\times 80$, resulting in an output impedance of 47.4Ω .

For timing analysis, we measure chiplet-to-chiplet communication delay and skew between all the wires in data bus as well as with the clock from end-to-end. We perform the timing analysis for our design by generating a transmission line model for the interposer interconnect channel. The interconnect lengths in our design varies from 200 to $9370 \mu\text{m}$ as shown in Fig. 15.

For a transmission line model, we generated a parameterized HSPICE model using the machine learning (ML)-based algorithm in [22]. In order to characterize the electrical properties of transmission line, a full-wave EM simulation with a large frequency range from dc to gigahertz region should be performed, which take a long CPU

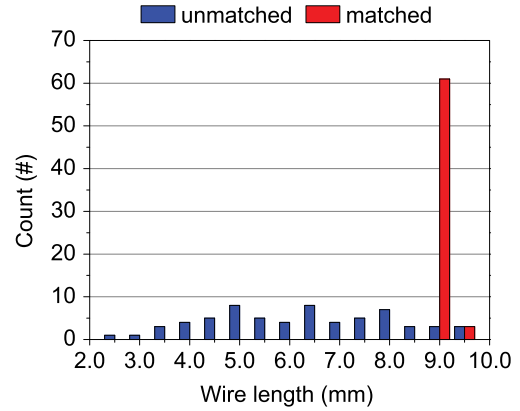


Fig. 13. Effect of MG in interposer routing result.

TABLE IV
2.5-D INTERPOSER DESIGN RESULTS [SEE FIG. 14(a)]

Routed net #	1,420
Metal layers used	4
Min wirelength	$200.0 \mu\text{m}$
Ave wirelength	$3568.9 \mu\text{m}$
Max wirelength	$9370.0 \mu\text{m}$
Via usage	5,328
PDN DC R	$12.2 m\Omega$
Area	116.64mm^2

time. Therefore, we generate the surrogate model of interposer transmission line using efficient Bayesian framework (EBF) [23], which replaces the EM solver to resolve this issue.

To create the surrogate model using a Gaussian process (GP), we first determine the samples based on uniform Latin hypercube sampling (LHS) and extract RLGC matrices at single frequency points instead of sweeping the full range. As each sample is characterized at a single frequency point, the total CPU time to collect training data is reduced significantly. The collected samples are then standardized to have zero mean and unit standard deviation, and used to perform the training of GP model to predict RLGC matrices of the transmission line in the interposer layer. Finally, we perform the propagation delay analysis of all the interconnect channels in the design by incorporating corresponding RLGC models into our HSPICE circuits.

The propagation delays and energy values for silicon interposer interconnects are shown in Fig. 16. We obtain the worst case propagation delay to be 104.50 ps . As our design is targeted to run at a frequency of 1 GHz , these longest propagation delay is well within the limits to meet the setup and hold times of the receiver. For $0.2\text{--}10.0\text{-mm}$ range, as wire becomes longer, both propagation delay and energy increase linearly, by 6.423 ps/mm and $0.037 \text{ pJ/bit} \times \text{mm}$, respectively.

In power analysis, we obtain each power of the chiplet core and AIB to estimate the total power of interposer system. Each routed net in interposer layer which is connected between two AIBs has the different wire length. However, this difference is not reflected in logic synthesis tool, so the power estimation

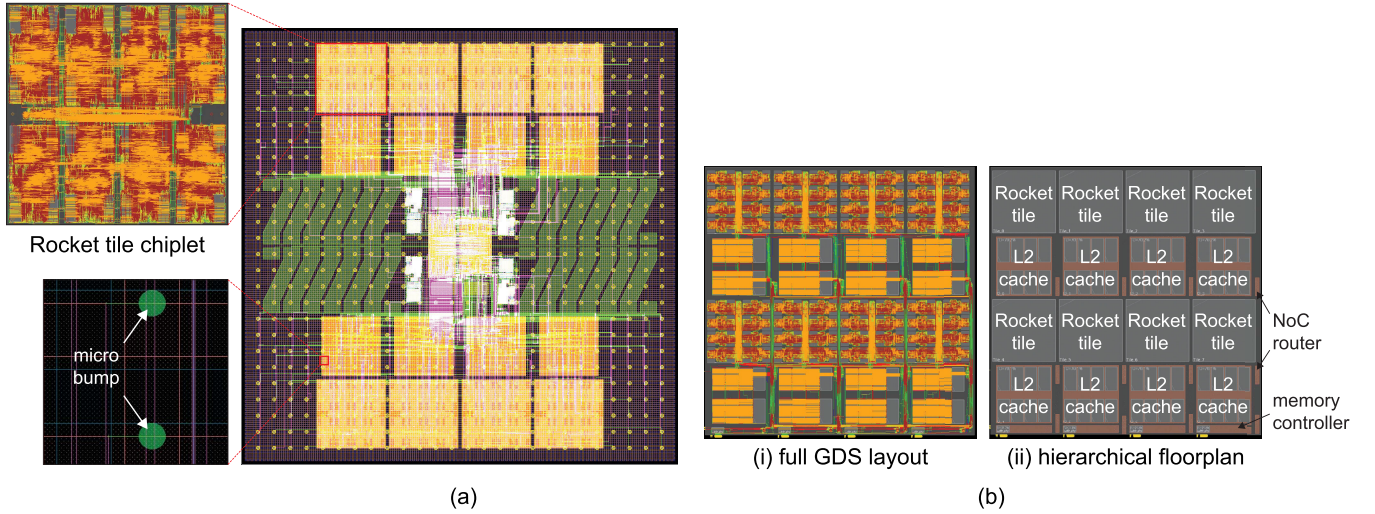


Fig. 14. (a) Interposer-based 2.5-D design versus (b) monolithic 2-D (GDS layouts) design (i) full GDS layout and (ii) hierarchical floorplan.

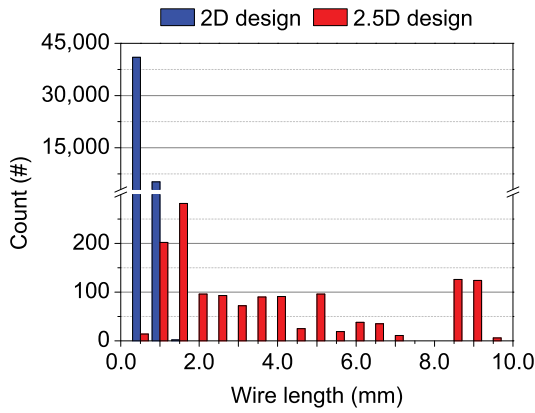


Fig. 15. Wire length distribution of 2.5-D design and comparison with 2-D counterpart.

result from the tool differ from the actual one. Moreover, the power loss occurs in IVR chiplets because the power of each chiplet is supplied through IVRs. Therefore, the power estimation in our EDA flow reflecting the wire length correctly is as follows:

$$P_{2.5D} = P_{CORE} + P_{I/O} + P_{PM} \quad (1)$$

where $P_{2.5-D}$ is total power of 2.5-D design, P_{CORE} is the power of chiplet cores, $P_{I/O}$ is the power of AIBs and P_{PM} is the power of power manage modules such as IVRs.

We run HSPICE simulation of the testbench with self-generated SPICE models for $P_{I/O}$. The power estimation of each chiplet core is done by Synopsys PrimeTime, and we obtain the power loss of IVRs from their power delivery efficiency.

C. Interposer SI and PI

1) *SI*: We perform the SI analysis and generated the eye diagram by converting the RLGC matrices of transmission line model into corresponding S-parameters and feeding them into

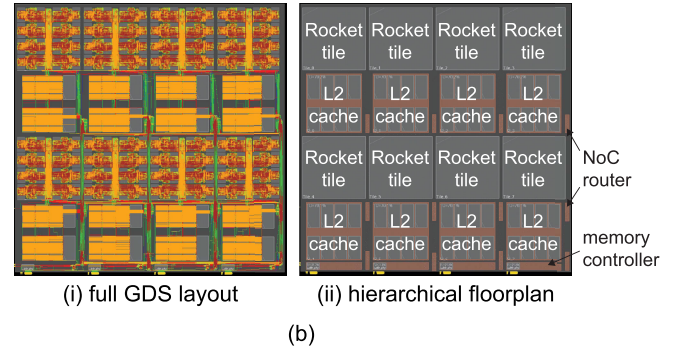


Fig. 16. (a) Propagation delay and (b) signal energy through interposer interconnections.

Keysight ADS. Our routing involves the use of complex interconnect structures as shown in Fig. 17(a) because they help in reducing the crosstalk compared to the simple structures. Therefore, we focus on a complex interconnect channel for crosstalk analysis. The characteristics of eye diagram are as follows: eye width is 0.995 ns, and eye height is 0.860 V. These results are obtained based on simulations done at the data rate of 1 Gb/s, I/O driver impedance of 50Ω considering ideal case and receiver chiplet pad parasitic of 2-pF capacitance.

2) *PI*: The PI of our design is ensured with the use of four IVR chiplets and distributing power through a mesh-type PDN as shown in Fig. 18. Since the clock frequency increases up to several gigahertz, modeling and analyzing interposer PDN require enormous computing resources since PDN mesh becomes a large structure. Therefore, we divide our mesh-type PDN into $M \times N$ unit cells using TMM [14] as shown in Fig. 19(a). Each unit cell is modeled as a lumped Π model which consist of R , L , G and C based on its physical and

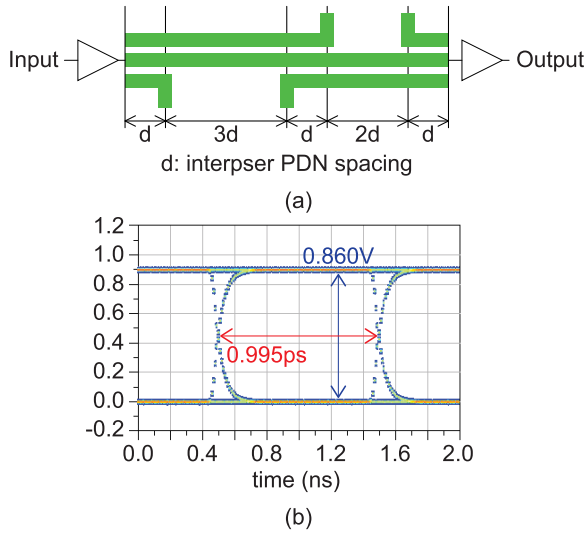


Fig. 17. Complex interconnect channel model and SI result in silicon interposer. (a) Complex interconnect channel model in the interposer. (b) Eye diagram of complex channel model.

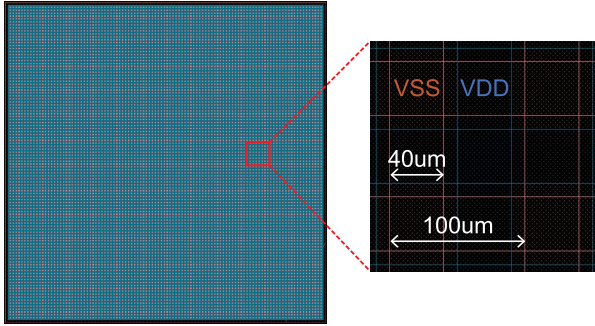


Fig. 18. PDN mesh in our interposer 2.5-D design.

electrical characteristics as shown in Fig. 19(b). They are obtained as follows [24]:

$$R = R_s \cdot \frac{S}{4W} \quad (2)$$

$$L = S \left[0.13e^{(-S/45)} + 0.14 \ln \left(\frac{S}{4W} \right) + 0.07 \right] \quad (3)$$

$$C_i = \frac{\epsilon_r}{10^3} \left[(44 - 28H)W^2 + (280 + 0.8S - 64)W \dots \right. \\ \left. \dots + 12S - 1500H + 1700 \right] \quad (4)$$

$$C_f = \epsilon_0 \epsilon_r 10^9 \left[\frac{4SW \left(\ln \frac{S}{S'} + e^{-1/3} \right)}{W\pi + 2H \left(\ln \left(\frac{S}{S'} + e^{-1/3} \right) \right)} + \frac{2S}{\pi} \sqrt{\frac{2H}{S'}} \right] \quad (5)$$

$$C = C_i + C_f \quad (6)$$

$$G = 2\pi \cdot f \cdot C \cdot \tan(\delta) \quad (7)$$

where R_s is surface resistance, W and S are the width/spacing of PDN mesh as shown in Fig. 19(b), $S' = S - 2W$, and H is the separation between P/G layer.

The results of PI analysis are shown in Fig. 20. Our interposer design shows PDN dc resistance of 12.2 m Ω and the first resonance peak at 1 GHz. Our IVR chiplet has dynamic voltage scaling speed (DVFS) of 1.65 V/ μ s, the settling time of 261 ns and the conversion efficiency of 79.0% with a

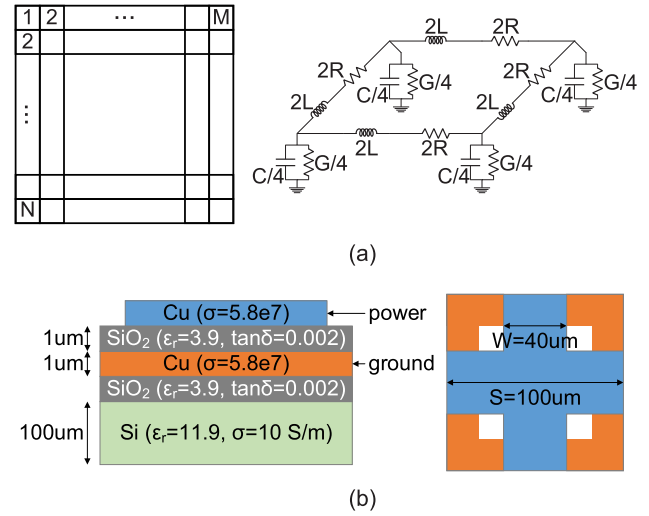


Fig. 19. Model of mesh-type PDN on silicon interposer used in our 2.5-D design. (a) Unit cell view. (b) Interposer PDN unit cell.

switching frequency (f_{sw}) of 125 MHz. In our 2.5-D design, the efficiency of IVR chiplet is limited by the inductor and capacitor technologies. We use the embedded inductors each of which inductance is 25 nH and low profile silicon capacitors each of which capacitance is 200 nF because of the limited area of silicon interposer. Therefore, f_{sw} has increased up to 125 MHz to reduce the output voltage ripple. The higher f_{sw} reduces the voltage settling time, however, it increases the switching loss and reduces the conversion efficiency in IVR chiplet.

VI. DSE RESULTS

A. Scalability of ROCKET-64

1) *Overhead of Chiplet Interface:* The 2.5-D integration requires chiplet interface modules in each chiplet design for the chiplet-to-chiplet communication. The additional module should be carefully designed not to compromise the performance of chiplet design. Therefore, we analyze the overhead of chiplet interface including protocol translator, SerDes and AIBs in our chiplet designs. Table V shows the proportion of chiplet interface in Rocket chiplet in terms of the cell count and power at 1 and 2 GHz.

At 1 GHz, the number of logic gate of chiplet interface is 0.78% of Rocket chiplet. Moreover, the chiplet interface only consumes 0.011 W, which is 1.05% of the total. Compared to the entire Rocket chiplet, the proportion of chiplet interface is negligible in both area and power consumption. We have also increased the operating frequency to 2 GHz to see the overhead of chiplet interface at the higher frequency.

At 2 GHz, the chiplet interface occupies 0.71% of the total gate count, and consumes 0.85% of the total chiplet power. Although the power consumption of chiplet interface has increased by $2.17 \times$ at 2 GHz when compared to 1 GHz, 0.85% of the total power consumption is still negligible. As our AIBs are overdesigned at 1 GHz to cover higher frequency, the cell count of AIBs remain the same in both frequencies. Moreover, as the translators and protocol FSMs for Hybrid-Link are

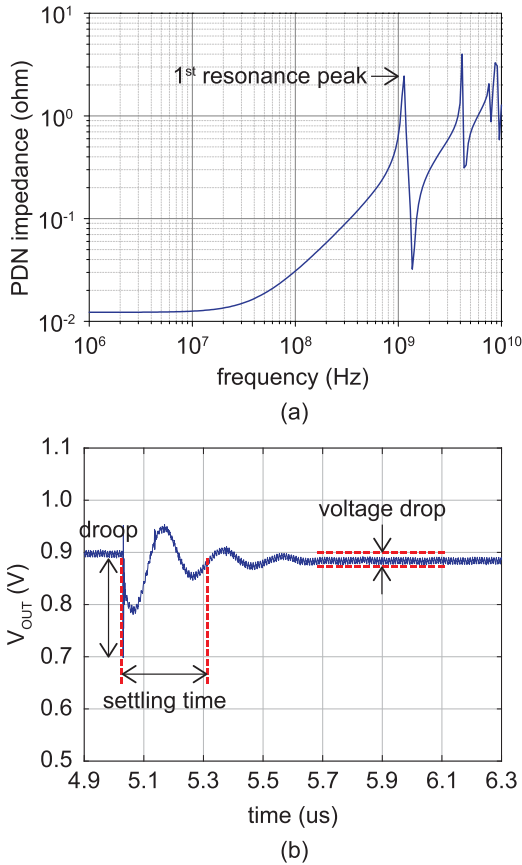


Fig. 20. PI results of our silicon interposer design. (a) Interposer PDN impedance. (b) Transient response of IVR chiplet.

TABLE V

AREA AND POWER IMPACTS OF CHIPLET INTERFACE IN ROCKET CHIPLET

Operating frequency		1GHz	2GHz	
Cell Count				
Rocket chiplet (#)		923,764	1,014,833	1.10×
Chiplet interface (#)	Total	7,226	7,213	1.00×
	Translator	5,960	5,949	1.00×
	SerDes	206	204	0.99×
	AIBs	1,060	1,060	1.00×
Interface/Rocket (%)		0.78	0.71	-0.07%
Power Consumption				
Rocket chiplet (W)		1.035	2.782	2.69×
Chiplet interface (W)	Total	0.011	0.024	2.17×
	Translator	0.005	0.011	2.38×
	SerDes	0.0004	0.0007	2.03×
	AIBs	0.006	0.012	2.02×
Interface/Rocket (%)		1.05	0.85	-0.20%

pipelined, the operating frequency of the chiplet is not limited by the additional interface modules. The tail latency is only affected due to the serialization of interfaces.

2) *Area Overhead Versus the Number of RISC-V Cores:* As chiplets are plug-and-play modules, we can increase the number of RISC-V cores with the extra area overhead. To increase the number of cores, we consider two approaches as follows:

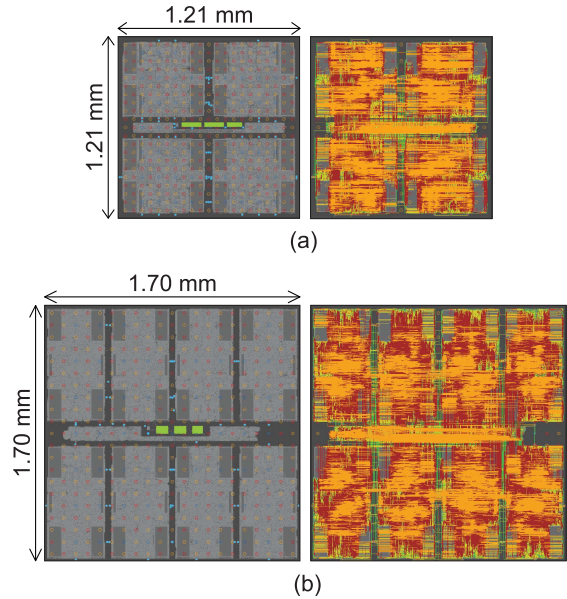


Fig. 21. Rocket chiplet designs with different number of cores. (a) Quad-core versus (b) octacore. AIBs are highlighted in blue and translators in green.

- 1) increasing the number of RISC-V cores in the Rocket chiplet itself;
- 2) increasing the number of Rocket chiplets in 2.5D design.

From the first approach, we have designed quad-core Rocket chiplet and compared to existing octacore Rocket chiplet to observe the area overhead as shown in Fig. 21. When the number of cores is doubled from four to eight, the area of Rocket chiplet has increased by 1.97×. As we have only increased the number of cores in Rocket chiplet, other chiplet designs remain the same. Therefore, the extra area overhead is the increased area of Rocket chiplet.

The number of RISC-V cores can be increased by adding Rocket chiplets in the design as well. However, in this approach, L2 cache chiplet is also needed for each additional Rocket chiplet because two different Rocket chiplets are not allowed to share one L2 cache chiplet in the current architecture. Besides, NoC chiplet should be redesigned to accept the additional Rocket chiplet with the higher I/O count. Therefore, the overall area overhead in this case comes from the area of additional Rocket and L2 cache chiplets, and the increased area of NoC chiplet. Comparing two different approaches, the first scenario is better than the second one to increase the number of cores in terms of the area overhead.

B. Monolithic 2-D Versus Interposer-Based 2.5-D

In monolithic 2-D design, we perform hierarchical design so that it has the same structure as interposer based 2.5-D design. We map all modules without I/O drivers and a power management module such as IVR on a single chip and split NoC module into 12 separate routers for the efficient configuration. We use the commercial 28 nm as the technology node and Cadence Innovus as the physical design tool. The layout and PPA analysis results of monolithic 2-D design with the target frequency of 1 GHz are shown in Fig. 14(b) and

TABLE VI
DESIGN COMPARISON BETWEEN MONOLITHIC 2-D AND INTERPOSER-BASED 2.5-D DESIGN USING FOUR IVR CHIPLETS

	2D Design	2.5D Design	
Area, Timing, Cell Count			
Area (mm^2)	53.14	116.64	2.19×
Footprint ($mm \times mm$)	7.29×7.29	10.80×10.80	-
Frequency (GHz)	1.0	1.0	-
Cell count (#)	8,047,741	7,552,762	0.94×
Routed Wirelength			
Min. wirelength (μm)	0.2	200.0	-
Avg. wirelength (μm)	222.4	3,568.9	16.05×
Max. wirelength (μm)	1435.1	9,370.0	6.53×
Power Consumption			
Total power (W)	8.984	11.574	1.29×
Logic power (W)	8.984	8.540	0.95×
I/O power (W)	-	0.256	-
IVR power (W)	-	2.778	-

Table VI. The total power is 8.948 W and the area of design including 64 RocketCores is 53.14 mm^2 .

In 2.5-D design, the overall area has increased by 2.19 \times compared to monolithic 2-D design. The main reason for the increase is the addition of power management modules including passive L and C because logic synthesis and P&R flow optimizes the logical area of chiplet. As the overall area of 2.5-D design has increased, the average length of routed wires in 2.5-D design is increased by 16.05 \times compared to 2-D design as shown in Table VI.

In power consumption, the total power of 2.5-D design has increased by 27.56% compared with 2-D design. The logic power in 2.5-D design is lower than monolithic 2-D design: 8.984 versus 8.392 W. This is because in 2-D design, the number of channels in the NoC module is higher than 2.5-D design, which causes the NoC module in 2-D to consume more power than the NoC chiplet in our 2.5-D design. As we do not use a package-based protocol in 2-D, it is necessary to increase the number of channels to handle additional traffic. However, AIBs are added for chiplet-to-chiplet communication and additional IVR chiplets cause the extra power loss in interposer-based 2.5-D design. Therefore, the overall power in 2.5-D design is higher than 2-D design.

C. Off-Chip VRM Versus On-Chip IVR

In 2.5-D chiplet integration, the optimal selection and analysis of power delivery configuration are fundamental to its performance and reliability. To investigate the tradeoffs between different power delivery schemes, we present two carefully designed on-board/on-interposer power delivery schemes for heterogeneous 2.5-D designs. As shown in Fig. 22, Design P1 uses a single on-board VRM to power up the entire 2.5-D IC, while Design P2 has four IVR chiplets which provide up to 12 A of current. Both VRM and IVR chiplet convert the external supply voltage which level is 3.6 V to the internal supply voltage of 0.9 V for 28-nm chiplets.

The comparisons between two 2.5-D designs in terms of PPA and PI are shown in Table VII. As Design P2 has four IVR chiplets, the number of interposer nets has increased by

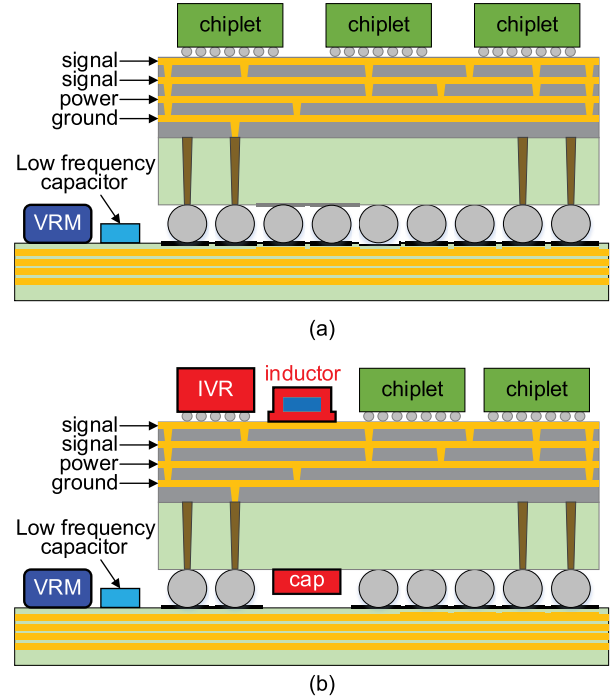


Fig. 22. Stack-up comparison of the two power delivery configurations: off-chip VRM versus IVR chiplet. (a) Design P1: off-chip VRM. (b) Design P2: IVR chiplet on the interposer.

1.03 \times compared to Design P1 due to the additional control signals. The area of Design P2 is 1.97 \times larger than Design P1 due to the additional IVR chiplets and passive components. As the area of design has increased, the maximum wirelength has also increased by 1.34 \times in Design P2. In terms of power consumption, logic powers are same in all designs while the I/O power has increased in Design P2 because it has longer interposer wires. The total power of Design P2 is 11.574 W, which is 1.22 \times higher than Design P1 due to the additional power loss of IVR chiplets.

As the current path for Design P1 includes the extra parasitic of P/G plane pairs on PCB, TSVs and C4 bumps compared with Design P2, Design P1 shows more resistive behavior with 15.9 m Ω of PDN dc resistance, compared to 12.2 m Ω of Design P2. Moreover, the first resonance peak in Design P1 comes at 600 MHz, whereas at 1 GHz in Design P2. This shows 1.67 \times bandwidth improvement in Design P2 approximately.

In Design P2, IVR chiplet with a switching frequency (f_{sw}) of 125 MHz significantly reduces the voltage settling time to 261 ns, compared to 23 μs in Design P1 with 2 MHz of f_{sw} . Moreover, dynamic voltage and frequency scaling (DVFS) is evaluated as 1.65 V/ μs in Design P2, compared to 0.18 V/ μs in Design P1. However, Design P1 shows 92.6% of voltage conversion efficiency, while Design P2 has 76.0%. In Design P2, the efficiency of IVR chiplet is primarily limited by the inductor and capacitor technologies. Due to the limited area of the silicon interposer, Design P2 uses inductors and capacitors which are 88 \times and 75 \times lower than Design P1. Therefore, the output voltage ripple only can be reduced by increasing f_{sw} up to 125 MHz. This in turn increases the switching

TABLE VII

COMPARISON OF PDN CONFIGURATIONS IN 2.5-D DESIGNS: OFF-CHIP VRM VERSUS ON-CHIP IVR

	Design P1	Design P2	
Power delivery config.	off-chip VRM	on-chip IVR	
Area, Routed Wirelength			
Area (mm^2)	59.29	116.64	1.97×
Routed net (#)	1,385	1,420	1.03×
Min. wirelength (μm)	630.0	200.0	0.32×
Avg. wirelength (μm)	2,807.9	3,568.9	1.27×
Max. wirelength (μm)	6,967.4	9,370.0	1.34×
Power Consumption			
Total power (W)	8.750	11.574	1.32×
Logic power (W)	8.540	8.540	1.00×
I/O power (W)	0.210	0.256	1.17×
IVR power (W)	-	2.778	-
Power Integrity			
VRM/IVR count (#)	1	4	-
Conversion Ratio (V/V)	3.6/0.9	3.6/0.9	-
PDN occupancy (%)	62.35	61.65	-0.70%
PDN DC R ($m\Omega$)	15.9	12.2	0.77×
1 st resonance freq. (GHz)	0.6	1.0	1.67×
Decap. (nF)	25	25	-
Inductor (L , nH)	2,200	25	0.01×
Capacitor (C , nF)	15,000	200	0.01×
f_{SW} (MHZ)	2	125	62.50×
Settling time (ns)	23,000	261	0.01×
DVFS ($V/\mu s$)	0.18	1.65	9.17×
Power efficiency (%)	92.6	76.0	-16.6%

loss and reduces the conversion efficiency by 16.6%. The comparison of Design P1 and P2 shows tradeoffs between the power, performance, and conversion efficiency depending on the power delivery configuration.

D. Silicon Versus Organic Interposers

A silicon interposer offers the best interconnect density with the RDL pitch less than 1 μm , however, it has a higher fabrication cost and poor channel characteristics compared to other interposer technologies. The organic interposer has been introduced as a promising alternative technology of the silicon interposer due to its low price and high-speed channel characteristics. However, the organic interposer has a limitation that its design rule is still larger than the silicon interposer despite the efforts to improve the organic interposer technology.

In this section, we choose liquid crystal polymer (LCP) as an organic substrate of the interposer and perform comparative analysis between our silicon and LCP designs to show the tradeoffs between PPA and reliability. Table VIII shows the design rules of silicon and LCP interposers which are used in these experiments. We choose LCP interposer technology which has 8- μm -pitch RDLs and 150- μm -pitch microbumps based on Panasonic R-F705S [25].

Table IX summarizes our design and analysis results of two 2.5-D IC designs using silicon and LCP interposers. The worst propagation delay of LCP interposer wire is smaller than silicon design due to the smaller resistance and capacitance of interposer wires. Even though the maximum wirelength is 1.80× higher, LCP design has 0.75× shorter delay than silicon design. The area of LCP design has increased by

TABLE VIII

DESIGN RULES OF SILICON AND LCP INTERPOSER TECHNOLOGIES USED IN THE COMPARATIVE EXPERIMENT

	Silicon	LCP (Organic)
Metal layer #	4	5
Metal thickness	1 μm	9 μm
Dielectric thickness	1 μm	25 μm
Dielectric constant	3.9	3.1
Min. line width/spacing	0.4 μm /0.4 μm	4 μm /4 μm
Via size	0.7 μm	6 μm
Through-via size/depth	10 μm /100 μm	40 μm /100 μm
Die-to-die spacing	100 μm	150 μm
micro bump pitch	40 μm	150 μm
C4 bump pitch	400 μm	800 μm
PDN width/spacing	80 μm /200 μm	80 μm /200 μm

TABLE IX

2.5-D IC DESIGN RESULTS COMPARISON: SILICON VERSUS LCP INTERPOSERS

	Silicon	LCP	
Timing, Area, Metal Layer Usage			
Frequency (GHz)	1.0	1.0	-
Interposer worst delay (ps)	174.90	131.66	0.75×
Area (mm^2)	116.64	466.56	4.00×
Interposer metal layer (#)	4	5	+1
Routed Wirelength in Interposer			
Min wirelength (mm)	0.14	0.43	2.99×
Avg wirelength (mm)	2.81	4.95	1.76×
Max wirelength (mm)	7.05	12.67	1.80×
Power Consumption			
Total power (W)	12.636	13.959	1.10×
Logic power (W)	8.540	9.312	1.09×
I/O power (W)	0.530	0.705	1.33×
IVR power (W)	3.566	3.942	1.11×
Signal Integrity			
Eye width (ns)	0.975	0.975	1.00×
Eye height (V)	0.816	0.637	0.78×
Power Integrity			
Interposer PDN occupancy (%)	61.65	61.65	1.00×
Interposer PDN DC R($m\Omega$)	17.24	10.08	0.59×
1 st resonance freq. (GHz)	0.50	1.17	2.34×
Output voltage ripple (mV)	12	16	1.33×
Initial ringing (V_{PP} , mV)	288	467	1.62×
Power efficiency (%)	71.78	71.76	-0.02%

4.00× when compared to silicon design due to the larger physical dimensions of LCP interposer technology as shown in Table VIII. Moreover, LCP design have used one additional metal layer to route all 1420 net on the interposer layer. As the area of LCP design has increased, the average length of interposer wires has also increased by 1.80×.

In LCP design, the total power is 13.959 W which is higher than silicon design by 10.46% with 0.02% reduction in power delivery efficiency as well as the increases in chiplet power and I/O power. The logic chiplet power has increased by 1.09× due to the larger sizes of chiplets. As the sizes of chiplets increase, the switching power has increased by 1.15× which led to the increase in overall chiplet power. The I/O power has also increased by 1.33× in LCP design because LCP interposer has longer interconnections than the silicon interposer.

In SI analysis, LCP design has the same eye width, but 21.94% smaller height when compared to silicon design. As the resistance of LCP interposer wire is smaller than silicon, the reflections at the receiver side play bigger role than silicon design. Therefore, the eye distortion in LCP design becomes worse due to the intersymbol interference (ISI).

In terms of PDN impedance, silicon design shows 17.24 m Ω of PDN dc impedance compared to 10.08 m Ω in LCP design. Moreover, the first resonance peak in silicon design comes at 0.50 GHz, whereas at 1.17 GHz in LCP design. It shows that LCP design has 2.34 \times better bandwidth than silicon design.

For the transient analysis of PI, the voltage settling time is 289 ns and DVFS is evaluated as 200 mV/439 ns in both silicon and LCP designs. However, LCP design shows 1.62 \times higher initial ringing and 1.33 \times larger output ripple at the output node of IVR chiplet due to lower L and C of interposer PDN. In terms of power delivery efficiency, silicon design shows 71.78%, while LCP design has 71.76%. The efficiency loss of 0.02% from the loss due to the higher output voltage ripple. The quantitative comparisons between silicon and LCP designs show the tradeoffs in PPA, SI and PI and provide the insight on the selection of interposer technology with given target design.

VII. CONCLUSION

In this article, we presented our vertically integrated EDA flow, which covers and fully automates the whole design phases of architecture, circuit and package. We verified our EDA flow by detailed descriptions of each step using a target design of ROCKET-64 with NoC configuration. We performed PPA comparison between 2.5-D IC and its monolithic 2-D counterpart to show the design overhead of interposer-based 2.5-D design. Moreover, we observed tradeoffs of power delivery schemes and interposer technologies with quantitative analyses in PPA, SI and PI. This work provides full sets of quantified comparison results of 2.5-D IC designs, which enables the SoC designer to have an objective criteria of evaluating the interposer-based design.

REFERENCES

- [1] G. E. Moore, "Cramming more components onto integrated circuits," *Proc. IEEE*, vol. 86, no. 1, pp. 82–85, Jan. 1998.
- [2] W. R. Davis *et al.*, "Demystifying 3D ICs: The pros and cons of going vertical," *IEEE Design Test Comput.*, vol. 22, no. 6, pp. 498–510, Jun. 2005.
- [3] J. Cong, G. Luo, J. Wei, and Y. Zhang, "Thermal-aware 3D IC placement via transformation," in *Proc. Asia South Pacific Design Automat. Conf.*, Jan. 2007, pp. 780–785.
- [4] K. Saban, "Xilinx stacked silicon interconnect technology delivers breakthrough FPGA capacity, bandwidth, and power efficiency," Xilinx, San Jose, CA, USA, White Paper Virtex-7 FPGAs, 2009.
- [5] T. Li, J. Hou, J. Yan, R. Liu, H. Yang, and Z. Sun, "Chiplet heterogeneous integration technology—Status and challenges," *Electronics*, vol. 9, no. 4, p. 670, 2020.
- [6] J. Kim *et al.*, "Architecture, chip, and package co-design flow for 2.5D IC design enabling heterogeneous IP reuse," in *Proc. 56th Annu. Design Automat. Conf. (DAC)*, Jun. 2019, pp. 1–6.
- [7] D. Stow, I. Akgun, R. Barnes, P. Gu, and Y. Xie, "Cost analysis and cost-driven IP reuse methodology for SoC design based on 2.5D/3D integration," in *Proc. 35th Int. Conf. Comput.-Aided Design (ICCAD)*, New York, NY, USA, 2016, pp. 56:1–56:6.
- [8] W. Liu, M.-S. Chang, and T. Wang, "Floorplanning and signal assignment for silicon interposer-based 3D ICs," in *Proc. 51st ACM/EDAC/IEEE Design Automat. Conf. (DAC)*, Jun. 2014, pp. 1–6.
- [9] M. A. Kabir and Y. Peng, "Chiplet-package co-design for 2.5D systems using standard ASIC CAD tools," in *Proc. 25th Asia South Pacific Design Automat. Conf. (ASP-DAC)*, Jan. 2020, pp. 351–356.
- [10] Y. Kim, J. Cho, K. Kim, V. Sundaram, R. Tummala, and J. Kim, "Signal and power integrity analysis in 2.5D integrated circuits (ICs) with glass, silicon and organic interposer," in *Proc. IEEE 65th Electron. Compon. Technol. Conf. (ECTC)*, May 2015, pp. 738–743.
- [11] H. Kalargaris and V. F. Pavlidis, "Interconnect design tradeoffs for silicon and glass interposers," in *Proc. IEEE 12th Int. New Circuits Syst. Conf. (NEWCAS)*, Jun. 2014, pp. 77–80.
- [12] K. Asanović *et al.*, "The rocket chip generator," Dept. EECS, Univ. California, Berkeley, Berkeley, CA, USA, Tech. Rep. UCB/EECS-2016-17, Apr. 2016. [Online]. Available: <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-17.html>
- [13] H. Kwon and T. Krishna, "OpenSMART: Single-cycle multi-hop NoC generator in BSV and chisel," in *Proc. IEEE Int. Symp. Perform. Anal. Syst. Softw. (ISPASS)*, Apr. 2017, pp. 195–204.
- [14] J.-H. Kim and M. Swaminathan, "Modeling of irregular shaped power distribution planes using transmission matrix method," *IEEE Trans. Adv. Packag.*, vol. 24, no. 3, pp. 334–346, Aug. 2001.
- [15] M. Lee *et al.*, "Automated generation of all-digital I/O library cells for system-in-package integration of multiple dies," in *Proc. IEEE 27th Conf. Electr. Perform. Electron. Packag. Syst. (EPEPS)*, Oct. 2018, pp. 65–67.
- [16] R. Chaware, K. Nagarajan, and S. Ramalingam, "Assembly and reliability challenges in 3D integration of 28 nm FPGA die on a large high density 65 nm passive interposer," in *Proc. IEEE 62nd Electron. Compon. Technol. Conf.*, May 2012, pp. 279–283.
- [17] D. Kehlet, *Accelerating Innovation Through a Standard Chiplet Interface: The Advanced Interface Bus (AIB)*. Accessed: Oct. 26, 2018. [Online]. Available: <https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/accelerating-innovation-through-aib-whitepaper.pdf>
- [18] V. C. K. Chekuri, N. Dasari, A. Singh, and S. Mukhopadhyay, "Automatic GDSII generator for on-chip voltage regulator for easy integration in digital SoCs," in *Proc. IEEE/ACM Int. Symp. Low Power Electron. Design (ISLPED)*, Jul. 2019, pp. 1–6.
- [19] H. M. Torun, M. Swaminathan, A. K. Davis, and M. L. F. Bellaredj, "A global Bayesian optimization algorithm and its application to integrated system design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 4, pp. 792–802, Apr. 2018.
- [20] H. M. Torun *et al.*, "A spectral convolutional net for co-optimization of integrated voltage regulators and embedded inductors," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2019, pp. 1–8.
- [21] H.-C. Lee, Y.-W. Chang, and P.-W. Lee, "Recent research development in flip-chip routing," in *Proc. Int. Conf. Comput.-Aided Design (ICCAD)*. Piscataway, NJ, USA: IEEE Press, 2010, pp. 404–410. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2133429.2133515>
- [22] H. M. Torun and M. Swaminathan, "High-dimensional global optimization method for high-frequency electronic design," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 6, pp. 2128–2142, Jun. 2019.
- [23] H. M. Torun, M. Larbi, and M. Swaminathan, "A Bayesian framework for optimizing interconnects in high-speed channels," in *Proc. IEEE MTT-S Int. Conf. Numer. Electromagn. Multiphys. Modeling Optim. (NEMO)*, Aug. 2018, pp. 1–4.
- [24] J. Kim *et al.*, "Chip-package hierarchical power distribution network modeling and analysis based on a segmentation method," *IEEE Trans. Adv. Packag.*, vol. 33, no. 3, pp. 647–659, Aug. 2010.
- [25] Panasonic Corporation, *Flexible Circuit Board Materials LCP Liquid Crystal Polymer FELIOS LCP*. Accessed: Jan. 17, 2017. [Online]. Available: <https://industrial.panasonic.com/www/products/electronic-materials/circuit-board-materials/felios/felioslep>