# FoldedHexaTorus: An Inter-Chiplet Interconnect Topology for Chiplet-based Systems using Organic and Glass Substrates

Patrick Iff
*ETH Zurich*
Zurich, Switzerland
patrick.iff@inf.ethz.ch

Maciej Besta
*ETH Zurich*
Zurich, Switzerland
maciej.besta@inf.ethz.ch

Torsten Hoefler
*ETH Zurich*
Zurich, Switzerland
torsten.hoefler@inf.ethz.ch

*Abstract*—**Chiplet-based systems are rapidly gaining traction in the market. Two packaging options for such systems are the established organic substrates and the emerging glass substrates. These substrates are used to implement the inter-chiplet interconnect (ICI), which is crucial for overall system performance. To guide the development of ICIs, we introduce three design principles for ICI network topologies on organic and glass substrates. Based on our design principles, we propose the novel *FoldedHexaTorus* network topology. Our evaluation shows that the *FoldedHexaTorus* achieves significantly higher throughput than state-of-the-art topologies while maintaining low latency.**

**Code: https://github.com/spcl/FoldedHexaTorus**

## I. INTRODUCTION

Technology scaling has fueled the ever-increasing performance per cost of processors and accelerators for a long time. However, since the 22 nm process, each transition to a scaled-down process has been accompanied by a surge in non-recurring (NR) cost of over 50% [1]. As a result, designing chips in cutting-edge processes is only economically viable at high production volumes. Chiplets promise a solution to this problem, as a single chiplet can be reused for multiple products, while the NR cost due to design and validation is incurred only once. Additional advantages of chiplets include improved yield (and hence lower cost) due to their smaller size compared to monolithic chips, and the option to integrate heterogeneous chiplets (built with different processes) in a single package.

Splitting a monolithic chip into multiple chiplets creates the need for a high-throughput inter-chiplet interconnect (ICI), which is crucial for communication-intensive workloads such as machine learning training and inference or scientific simulations. The ICI is built using die-to-die (D2D) links [2], [3], [4], which are implemented on organic or glass substrates [5], [6], as well as silicon interposers [7] or bridges [8], [9]. While silicon interposers and bridges offer higher bandwidth, they come with higher production costs. Therefore, our work focuses on organic and glass substrates. Another advantage of these substrates over silicon interposers is that, since they use a different fabrication process, they are not bound by the reticle limit and thus allow the construction of massive systems [10].

A major determinant of ICI throughput is the topology of links between chiplets. For systems based on passive silicon interposers or silicon bridges, the ICI topology is restricted to connecting only adjacent chiplets, resulting in topologies such as *Mesh* and *HexaMesh* [11]. On active silicon interposers, the link length is unrestricted, and many topologies have been proposed [12], [7], [13], [14], [15]. For organic and glass substrates, the link length is less restricted than on passive interposers (due to superior loss characteristics), but more restricted than on active silicon interposers (due to the absence of repeaters), opening up a new and largely unexplored design space for ICI topologies.

In this work, we develop design principles for ICI topologies on organic and glass substrates (**contribution 1**). These design principles reveal that, to achieve high throughput, the ICI topology must have a *low network radix*, a *low network diameter*, and *short links*—three properties that are inherently in conflict with one another [16]. By searching for a sweet spot in this design space, we conceive the novel *FoldedHexaTorus* topology (**contribution 2**). The *FoldedHexaTorus* features a constant network radix of six, a constant link length only slightly longer than the chiplet side, and a network diameter of less than $\sqrt{N}$, where $N$ is the total number of chiplets. Our evaluation (**contribution 3**) shows that, for chiplet-based systems with organic and glass substrates, *FoldedHexaTorus* outperforms topologies for passive and active silicon interposers, as well as network-on-chip (NoC) topologies.

## II. BACKGROUND

### A. Overview of Packaging Technologies

**Organic substrates** (see Fig. 1a) are a proven packaging technology with established supply chains. They are not bound by the lithographic reticle limit and enable the assembly of systems that surpass the size of monolithic or silicon-interposer-based chips. However, the large pitch of controlled collapse chip connection (C4) bumps limits the bandwidth of D2D links, making them a bottleneck.

**Glass substrates** (see Fig. 1b) are a packaging technology currently under development [17]. Compared to organic sub-
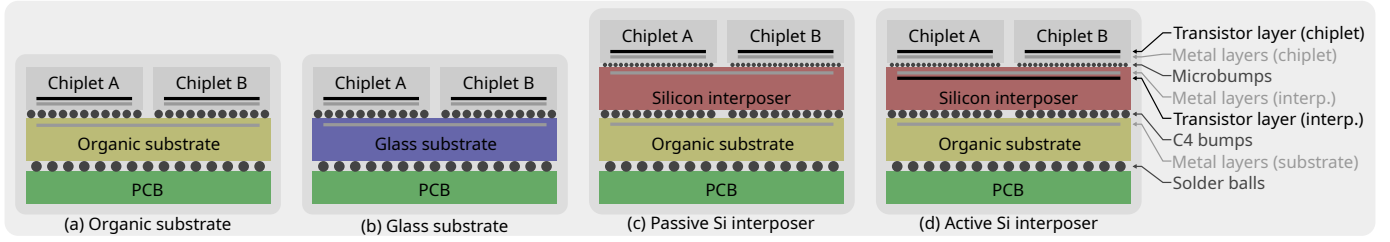
Fig. 1: **(§II-A) Overview of packaging technologies.**

strates, they promise smaller wire and bump pitches, superior thermal stability, and better electrical performance [5], [6].

**Passive silicon interposers** (see Fig. 1c) provide higher D2D link bandwidth than organic and glass substrates, enabled through the use of fine-pitch microbumps. However, passive silicon interposers increase manufacturing costs and complexity, and they suffer from severely limited link length [18].

**Active silicon interposers** (see Fig. 1d) alleviate link-length restrictions by providing a transistor layer, enabling the construction of repeaters and buffers [19]. However, this transistor layer further increases manufacturing cost and complexity, and can lead to thermal issues.

### B. Data Rate and Link Length

For organic and glass substrates, as well as passive silicon interposers, there is a trade-off between link length and data rate. Simulations using the transmission line model [20], as performed by Kim [21], show that as link length increases, the maximum admissible data rate decreases (see Fig. 2). For passive silicon interposers, the data rate drops significantly when the link length exceeds 4 mm. For organic and glass substrates, the decline in data rate is less severe and begins only at link lengths of 10–20 mm.

## III. DESIGN PRINCIPLES FOR ICI TOPOLOGIES

While topologies for data centers [22] and NoCs [23], [24] are often constructed following graph-theory-based design principles, such principles appear to be lacking for ICI topologies on organic and glass substrates. To fill this gap, we introduce three design principles for high-throughput ICI topologies on organic and glass substrates.

### A. Principle 1: Minimize the Network Diameter

Minimizing the network diameter (the maximum number of router-to-router hops per packet) has been a major design goal for many topologies in both the data center [25], [22] and NoC [23], [24] domains. The primary motivation for reducing the diameter is that fewer hops per packet translate into fewer packets processed by each router, thereby reducing congestion. Recall that in ICIs based on organic and glass substrates, the limited number of bumps connecting each chiplet to the substrate constitutes a bottleneck. While reducing the network diameter is also beneficial for ICI topologies on silicon interposers, it is even more critical for organic and glass substrates, as it eases pressure on the aforementioned bottleneck and increases throughput. Additionally, a smaller diameter results in lower latency, as each chiplet-to-chiplet hop

incurs time-consuming processing by physical layers (PHYs) and routers. Furthermore, since the energy consumption of a D2D link mainly depends on the number of bits transmitted, reducing the network diameter also decreases the overall energy consumption.

### B. Principle 2: Tune the Link-range

*Definition*: The **link-range** is the number of intermediate chiplets that a link stretches across (a link between adjacent chiplets has a range of zero).

As discussed in Section II-B, the maximum data rate decreases with increasing link length; hence, shorter links are preferred. However, lowering the network diameter (Principle 1) requires a higher link-range, raising the following question:

> *What link-range should be allowed for the best trade-off between a low network diameter and short links?*

The answer to this question depends on the packaging technology and the size of the chiplets. For passive silicon interposers, the link length is so restricted that only a link-range of zero is practical, leading to the use of topologies such as *Mesh* or *HexaMesh* [11]. For organic and glass substrates, the link length is less restricted, which motivates us to explore the use of higher link-ranges. For our analysis, we assume square chiplets of $74 \text{ mm}^2$ (the same area as in AMD's EPYC and Ryzen processors [26]), but we show in Section IV that our results generalize well when varying the chiplet size.
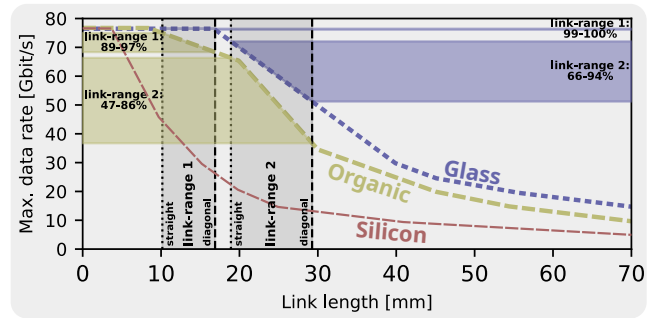


Fig. 2: **(§III-B) Relation between data rate and link length** based on simulations by Kim [21]. Yellow and blue areas show the achievable fraction of the max. data rate for a link-range of one and two on organic and glass substrates.

Fig. 2 displays the relationship between link length and data rate based on transmission line simulations performed by Kim [21]. We present a range of possible link lengths (the gray area), spanning from a perfectly straight link (the shortest) to a

diagonal link at a 45° angle (the longest) for link ranges of one and two. For both link-ranges, we show the percentage of the maximum data rate achievable in organic and glass substrates (the yellow and blue areas, respectively). We observe that for a link-range of one, links in glass substrates can operate at 99–100% of the maximum data rate, while the data rate in organic substrates drops to 89–97%. For a link-range of two, the data rate drops to 66% for glass and 47% for organic substrates in the worst case. We conclude that for chiplets of approximately $74\,\mathrm{mm}^2$, a link-range of one provides the best trade-off between low network diameter and short links.

## C. Principle 3: Minimize the Network Radix

The higher the network radix (number of D2D links per chiplet), the greater the area overhead of the chiplet, as each D2D link requires a dedicated PHY. Furthermore, recall that the number of C4 bumps per chiplet is limited, especially for organic and glass substrates. Approximately 50% of these bumps are used for the chiplet's power supply [5]; the remaining bumps are available for D2D links and off-chip I/O. As the network radix increases, the number of bumps per link—and therefore the per-link bandwidth—decreases. Since each link requires a constant number of bumps for non-data wires such as clock or handshake signals (12 for universal chiplet interconnect express (UCIe) [27]), the total off-chiplet bandwidth also decreases with increasing radix. These arguments suggest that the network radix should be minimized. However, lowering the network diameter (Principle 1) requires a higher network radix [16], creating a conflict between Principles 1 and 3 and raising the following question:

*What is the optimal balance between minimizing network diameter and network radix?*

This question is difficult to answer, as the exact relationship between radix and diameter is still an open research problem. While the Moore bound [16] provides a theoretical lower bound on the achievable diameter for a given radix and chiplet count, it cannot be directly applied to our setting: 1) for most combinations of chiplet count and network radix, it is unknown whether the Moore bound can actually be achieved, and 2) the Moore bound assumes links of arbitrary length, whereas we are constrained to links with range one (Principle 2). In Section IV, we address this question experimentally by evaluating different topologies with varying network radices.

## IV. PRINCIPLED DESIGN OF ICI TOPOLOGIES

We follow our design principles (see Section III) to construct ICI topologies for chiplet-based systems using organic or glass substrates. Consider a *Mesh* topology (see Fig. 3a), which is a common choice for chiplet-based systems (e.g., Tesla Dojo [28]). We leverage a link-range of one (Principle 2) to transform the *Mesh* into a *FoldedTorus* [29] (see Fig. 3d), which significantly lowers the network diameter (Principle 1). With a low network radix of four, *FoldedTorus* also complies with Principle 3; however, it remains unclear whether further reducing the network diameter at the cost of a higher network radix would improve performance. To answer this question, we apply our design principles to two additional topologies. *HexaMesh* [11] is a radix-6 topology (see Fig. 3b) with a lower network diameter than a *Mesh*. Again, we use links with range one (Principle 2) to reduce the network diameter, yielding a novel topology that we call *FoldedHexaTorus* (see Fig. 3e). As an example of a radix-8 topology, we consider a mesh-like topology with additional diagonal links (see Fig. 3c), named *OctaMesh*. By applying design Principles 1 and 2, we transform it into the *FoldedOctaTorus* (see Fig. 3f).

To find the best trade-off between network diameter and network radix, we evaluate the aforementioned topologies using the same methodology as in the main evaluation (see Section V). Recall that our analysis motivating the choice of a link range of one (Principle 2) was based on chiplets with an area of $74\,\mathrm{mm}^2$. To assess the robustness of our topologies across varying chiplet sizes, we repeat the simulations for chiplets with areas of $37\,\mathrm{mm}^2$ (half) and $148\,\mathrm{mm}^2$ (double).

Fig. 4 shows the throughput and latency for chiplet counts ranging from 16 to 256 on organic substrates. The results for glass substrates are similar, omitted due to space constraints, and can be found in our open-source repository. For almost all chiplet sizes and chiplet counts considered, *FoldedHexaTorus* achieves the highest throughput while providing the second-lowest latency. While throughput depends on various factors such as network radix, network diameter, link lengths (and consequently the maximum data rate), and others, latency is primarily determined by the network diameter.

Tables I and II present the area and power overhead of each topology relative to a *Mesh* topology on organic substrates (results for glass substrates are similar and omitted due to space constraints, they can be found in our open-source repository). The area overhead results from additional
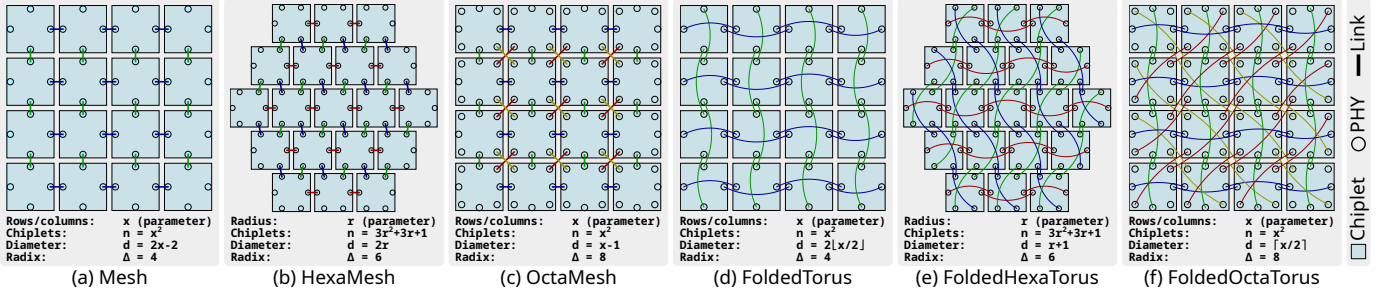


Fig. 3: **(§IV) Basic ICI topologies (a-c) and versions optimized for organic or glass substrates (d-f).**
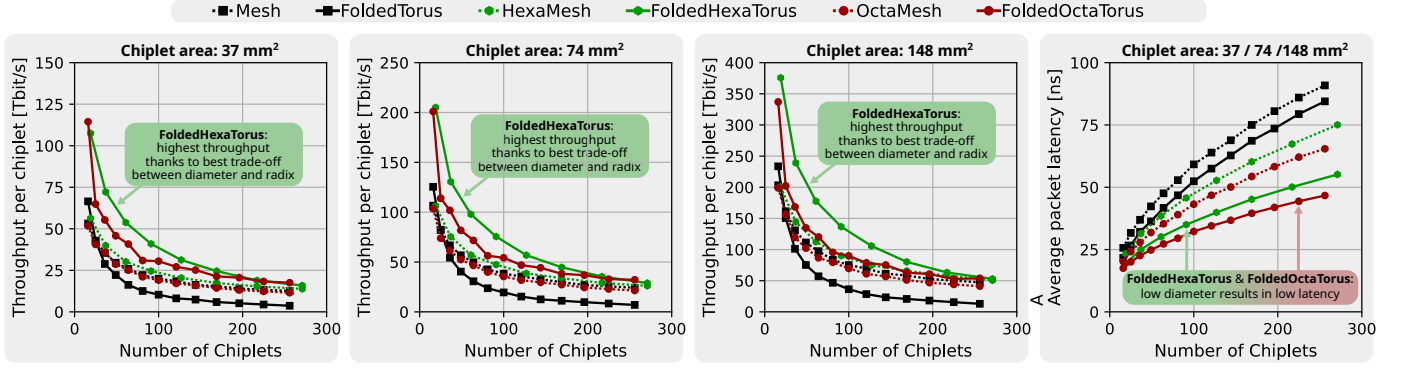
Fig. 4: (§IV) **Throughput and latency of ICI topologes.**

PHYs and thus depends only on the network radix and the chiplet size. In contrast, the power overhead (or, in many cases, power savings) relative to a *Mesh* topology depends on the number of bits transmitted per second over the links. As a result, it is influenced by both the network diameter (a smaller diameter means each packet traverses fewer links) and the maximum throughput (a topology that achieves higher throughput transmits more packets per second, thereby consuming more power). Thus, the power consumption of topologies optimized for ICIs on organic and glass substrates reflects a trade-off: power savings due to reduced diameter and power overhead due to increased saturation throughput.

| Topology | Chiplet area relative to a Mesh topology | | |
|---|---|---|---|
| | $37mm^2$ | $74mm^2$ | $148mm^2$ |
| Mesh | $0.00 \pm 0$ % | $0.00 \pm 0$ % | $0.00 \pm 0$ % |
| FoldedTorus | $0.00 \pm 0$ % | $0.00 \pm 0$ % | $0.00 \pm 0$ % |
| HexaMesh | $4.34 \pm 0$ % | $2.27 \pm 0$ % | $1.16 \pm 0$ % |
| FoldedHexaTorus | $4.34 \pm 0$ % | $2.27 \pm 0$ % | $1.16 \pm 0$ % |
| OctaMesh | $8.69 \pm 0$ % | $4.54 \pm 0$ % | $2.32 \pm 0$ % |
| FoldedOctaTorus | $8.69 \pm 0$ % | $4.54 \pm 0$ % | $2.32 \pm 0$ % |

TABLE I: **(§IV) Total chiplet area (including PHYs) relative to a *Mesh* topology** (mean over all chiplet counts).

| Topology | Power consumption relative to a Mesh topology | | |
|---|---|---|---|
| | $37mm^2$ | $74mm^2$ | $148mm^2$ |
| Mesh | $0.00 \pm 0.00$ % | $0.00 \pm 0.00$ % | $0.00 \pm 0.00$ % |
| FoldedTorus | $-0.81 \pm 0.58$ % | $-1.67 \pm 0.79$ % | $-3.40 \pm 1.50$ % |
| HexaMesh | $-0.12 \pm 0.06$ % | $-0.35 \pm 0.36$ % | $-0.74 \pm 0.74$ % |
| FoldedHexaTorus | $1.19 \pm 1.96$ % | $1.84 \pm 3.21$ % | $2.35 \pm 4.64$ % |
| OctaMesh | $-0.83 \pm 0.73$ % | $-1.69 \pm 1.40$ % | $-2.93 \pm 2.06$ % |
| FoldedOctaTorus | $0.28 \pm 1.73$ % | $-0.10 \pm 2.11$ % | $-1.66 \pm 2.37$ % |

TABLE II: **(§IV) Power at saturation throughput relative to a *Mesh* topology** (mean over all chiplet counts).

Due to its superior throughput, second-lowest latency, and moderate area and power overhead, we recommend *Folded-HexaTorus* as the topology of choice for chiplet-based systems on organic and glass substrates.

## V. EVALUATION

We compare the throughput, latency, area, and power consumption of our proposed *FoldedHexaTorus* topology against several baseline topologies across a wide range of system sizes. Our analysis spans organic and glass substrates, various architectures (homogeneous and heterogeneous chiplets), and both synthetic traffic patterns and real-world traces.

### A. Baseline Topologies

Table III lists the baseline topologies we compare against, and Fig. 5 visualizes a selection of them. While *Mesh* is commonly used in practice and *HexaMesh* has been proposed for both organic substrate- and silicon interposer-based systems, we are not aware of any topologies specifically designed for organic or glass substrates. Therefore, our comparison includes a broad spectrum of topologies originally proposed for silicon interposers [12], [7], [13], [14], [15], NoCs in monolithic chips [29], [30], [31], and computer networks [32]. Since some interposer topologies were originally designed for slightly different architectures, we adapt them to our setting—for example, by using on-chiplet instead of on-interposer routers. As these topologies are not optimized for organic and glass substrates, most violate at least one design principle: they either feature an unsuitable network radix (Principle 3), an excessive link range (Principle 2), or a suboptimal network diameter (Principle 1). *FoldedHexaTorus* is the only topology aligned with all three design principles.

| Topology | Diameter | Radix | Link-range |
|---|---|---|---|
| Mesh | $2\sqrt{N}-2$ | $4$ | $0$ |
| Torus | $2\lfloor\sqrt{N}/2\rfloor$ | $4$ | $\sqrt{N}-2$ |
| HexaMesh [11] | $\frac{\sqrt{12N-3}}{3}-1$ | $6$ | $0$ |
| DoubleButterfly[12] | $\sqrt{N}$ | $4$ | $\frac{\sqrt{N}}{2}-1$ |
| ButterDonut [7] | $\approx \lfloor\frac{2}{3}\sqrt{N}\rfloor$ | $4$ | $\frac{\sqrt{N}}{2}-1$ |
| ClusCross V1 [13] | $\sqrt{N}-1$ | $4$ | $\sqrt{N}-2$ |
| ClusCross V2 [13] | $\lceil\frac{3\sqrt{N}}{4}\rceil$ | $4$ | $\sqrt{N}-2$ |
| Kite Small [14] | $\sqrt{N}-1$ | $4$ | $0$ |
| Kite Medium [14] | $\sqrt{N}$ | $4$ | $1$ |
| Kite Large [14] | $\approx \sqrt{N}$ | $4$ | $1$ |
| SID-Mesh [15] | $\sqrt{N}-1$ | $4$ | $0$ |
| FoldedTrous [29] | $2\lfloor\sqrt{N}/2\rfloor$ | $4$ | $1$ |
| Hypercube [30] | $\log_2(N)$ | $\log_2(N)$ | $\frac{\sqrt{N}}{2}-1$ |
| FlattenedButterfly [31] | $2$ | $2\sqrt{N}-2$ | $\sqrt{N}-2$ |
| HoneycombMesh [32] | $1.63\sqrt{N}$ | $3$ | $0$ |
| HoneycombTorus [32] | $0.81\sqrt{N}$ | $3$ | $3\sqrt{N/6}-2$ |
| **FoldedHexaTorus** | $\frac{\sqrt{12N-3}}{6}+\frac{1}{2}$ | $6$ | $1$ |

TABLE III: **(§V-A) Evaluated topologies**. We highlight diameter, radix, and link-range, to indicate high , moderate , or low compliance with design principles.
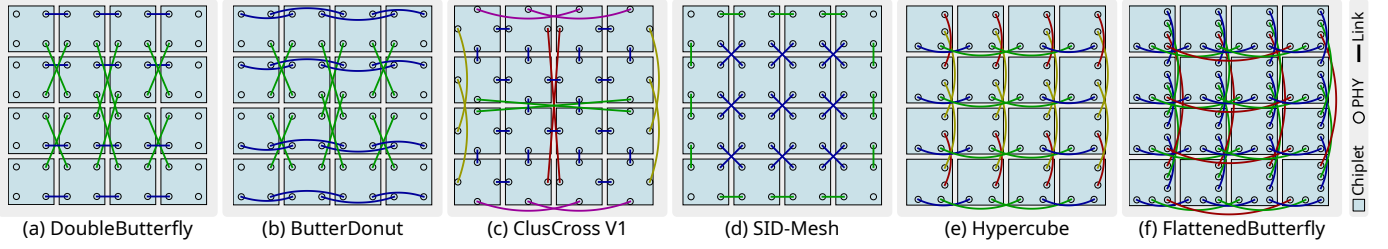
Fig. 5: **(§V-A) A selection of baseline topologies; proposed for silicon interposers (a-d) or NoCs (e-f)**.

## B. Evaluation Methodology

We measure saturation throughput and average packet latency using the cycle-based BookSim simulator [33], which models input-queued, pipelined routers. We use four virtual channels with 4-flit buffers each. We implement a custom routing algorithm based on Dijkstra's algorithm, incorporating the turn model [34], the simple cycle breaking algorithm [35], and a dual graph construction [36] to enable deadlock-free, shortest-path routing on arbitrary topologies. Table IV lists the remaining simulation parameters. We extend BookSim to support traffic trace simulation by integrating Netrace [37], [38]. To estimate the area overhead of PHYs and power consumption, we use the RapidChiplet toolchain [39].

| | Parameter | Organic | Glass | Reference |
|---|---|---|---|---|
| $S_c$ | Chiplet spacing | 150 $\mu m$ | 100 $\mu m$ | [5] (Table 1) |
| $A_c$ | Chiplet area | 74 mm$^2$ | 74 mm$^2$ | [26] (Page 6) |
| $A_p$ | PHY area | 0.88 mm$^2$ | 0.88 mm$^2$ | [27] (Tab. 29) |
| $P_c$ | Chiplet power | 25W | 25W | Assumption |
| $E_{bit}$ | Energy per bit | 0.3 pJ | 0.3 pJ | [2] (Page 1) |
| $L_p$ | PHY latency | 2ns | 2ns | [27] (Table 6) |
| $L_r$ | Router latency | 3ns | 3ns | Assumption |
| $f_{pb}$ | Bumps for power | 50% | 50% | [5] (Page 3) |
| $f_{io}$ | Bumps for off-chip I/O | 20% | 20% | Assumption |
| $N_c$ | Cores per chiplet | 8 | 8 | [26] (Page 6) |
| $P_b$ | Bump pitch | 50 $\mu m$ | 35 $\mu m$ | [5] (Table 1) |
| $N_w$ | Non-data wires | 12 | 12 | [27] (Fig. 73) |
| $\epsilon_r$ | Dielectric constant | 3.1 | 3.3 | [5] (Table 1) |
| $c$ | Speed of light | 299,792 km/s | 299,792 km/s | Constant |

TABLE IV: **(§IV) Parameters used in our experiments**.

*1) Throughput:* BookSim reports the relative throughput $T_r$, defined as the maximum rate at which each core can inject traffic into the network. We compute the absolute per-chiplet throughput $T_a$ as follows:

$$T_a = N_c \cdot T_r \cdot \left( \left( \frac{\overbrace{(A_c + R \cdot A_p)}^{\text{total chiplet area}} \cdot (1 - f_{pb} - f_{io})}{\underbrace{R \cdot P_b^2}_{\text{bumps per link}}} - N_w \right) \cdot \underbrace{\text{rate}\left(\widehat{L}\right)}_{\text{absolute per-core throughput}} \right)$$

with *data wires per link* over the numerator.

Here, $R$ denotes the router radix, and $\widehat{L}$ represents the maximum link length, computed via RapidChiplet, accounting for both the chiplet spacing $S_c$ and the physical location of a PHY within the chiplet. The function *rate*() returns the maximum achievable data rate for a given link length, as defined in Fig. 2. All remaining parameters are listed in Table IV.

*2) Latency:* We assume that each chiplet contains a router with latency $L_r$, which can relay messages either to PHYs (with latency $L_p$) or to cores. The latency $L_l$ of a link of length $L$ is modeled using the transmission line equation: $L_l =$

$L \cdot \sqrt{\epsilon_r}/c$, where $\epsilon_r$ is the relative permittivity of the medium and $c$ is the speed of light in vacuum. All parameters are listed in Table IV. Since BookSim is cycle-based, we set the cycle time to $1$, ns and configure all inputs accordingly. The computed link latency $L_l$ is rounded up to the next full cycle.

*3) Area:* The logic of a chiplet occupies an area of $A_c$, and each PHY contributes an additional area of $A_p$. Thus, the total area $A$ of a radix-$R$ chiplet is given by $A = A_c + R \cdot A_p$.

*4) Power:* The logic of a chiplet consumes $P_c$ watts. To estimate the power consumption of each PHY, we count the number of bits transmitted per second during BookSim simulations and multiply it by the energy per bit $E_{bit}$. Note that power values are reported at the highest possible throughput.
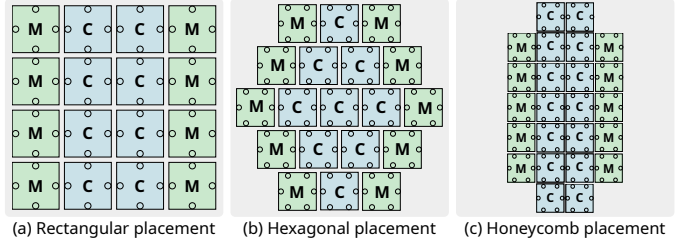


Fig. 6: **(§V-C) Placement of compute chiplets (C) and memory chiplets (M) in the heterogeneous architecture**.

## C. Results on Synthetic Random Uniform Traffic

We conduct a broad evaluation using random uniform traffic, as it is as a good proxy for many real-world applications such as graph computations, sparse linear algebra solvers, and adaptive mesh refinement [22]. We evaluate *FoldedHexaTorus* on two architectures: a homogeneous configuration with compute chiplets only, and a heterogeneous configuration with both compute and memory chiplets. In the heterogeneous case, memory chiplets occupy the leftmost and rightmost columns, while the remaining chiplets host compute cores (see Fig. 6). We use a 50/50 mix of core-to-core and core-to-memory traffic, following common practice [7], [14].

Fig. 7 compares the throughput, latency, area, and power of *FoldedHexaTorus* against the baseline topologies (see Section V-A) for both architectures and for both organic and glass substrates. We observe that *FoldedHexaTorus* achieves high throughput across all architectures, substrates, and chiplet counts. Notably, for *ClusCross*, *Torus*, *HoneycombTorus*, and *FlattenedButterfly*, throughput drops to zero once the system size exceeds a certain threshold, as some links surpass the maximum permissible length of $70$ mm. *FoldedHexaTorus* also
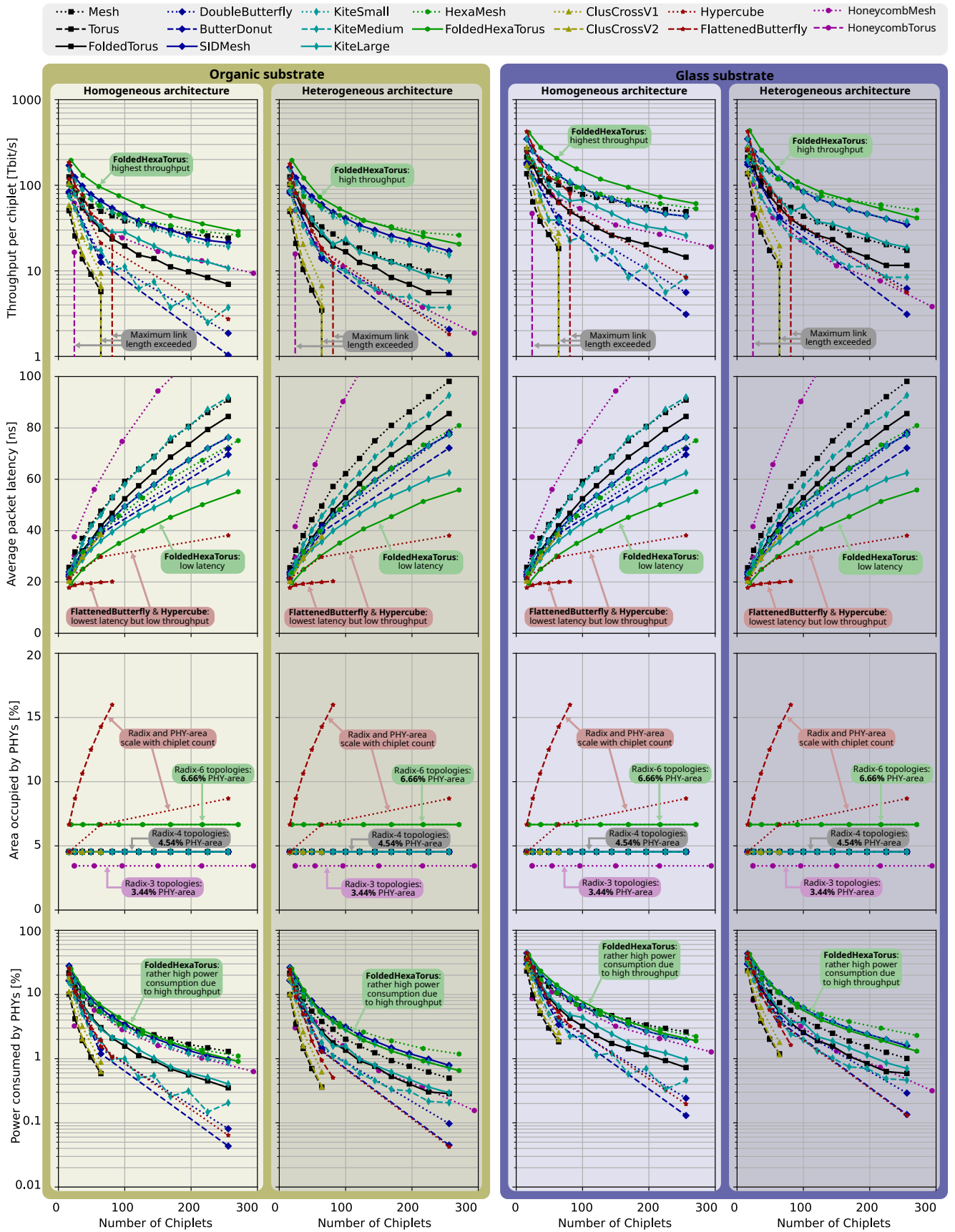
Fig. 7: (§V-C) **Throughput, latency, area, and power of ICI topologies** for varying chiplet counts (random uniform traffic).
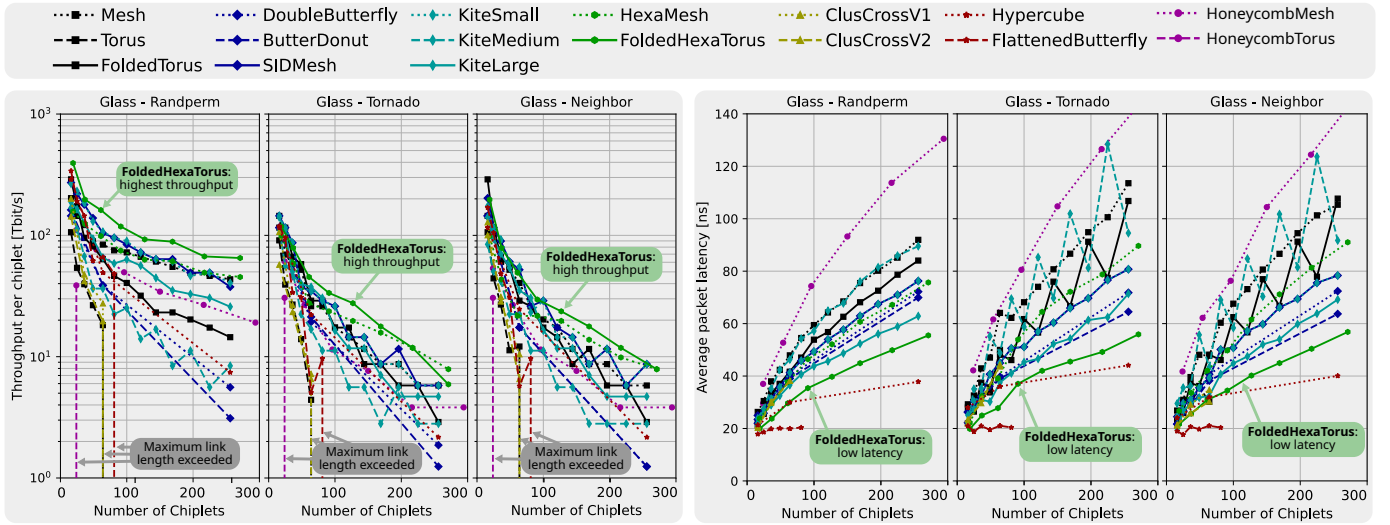
Fig. 8: **(§V-D) Throughput and latency of ICI topologies for different chiplet counts and traffic patterns.**

demonstrates excellent latency, only outperformed by *FlattenedButterfly* and *Hypercube*, both of which suffer from severely limited throughput.

The percentage of total silicon area occupied by PHYs depends solely on the network radix. Most topologies considered use a radix of four, resulting in a PHY area of $4.54\%$. *FoldedHexaTorus* and *HexaMesh* use a radix of six, slightly increasing the PHY area to $6.66\%$. The only topologies where the radix scales with the number of chiplets, rather than remaining constant, are *FlattenedButterfly* and *Hypercube*. Since PHY power consumption is proportional to the number of transmitted bits, it follows the same trend as throughput. As *FoldedHexaTorus* achieves the highest throughput, it also consumes a significant amount of power. However, despite this high throughput, *FoldedHexaTorus*'s power consumption remains comparable to several topologies with lower throughput.

### D. Results on Additional Synthetic Traffic Patterns

Fig. 8 shows the throughput and latency of *FoldedHexaTorus* and the baseline topologies under three additional synthetic traffic patterns: *Random Permutation*, *Tornado*, and *Neighbor*. We present results for the homogeneous architecture on a glass substrate; results for the organic substrate are similar and available in our repository. Performance trends align with those observed under random uniform traffic, with *FoldedHexaTorus* consistently achieving high throughput and low latency across all patterns.

### E. Results on Real-World Traffic Traces

We evaluate *FoldedHexaTorus* using the *blackscholes* and *fluidanimate* traces from the Netrace collection [40], based on the PARSEC benchmark suite [41]. Each trace is divided into five regions. Due to the traces spanning billions of cycles, simulating them in a cycle-accurate simulator is prohibitively time-consuming. Therefore, we simulate the first $100{,}000$ cycles of each region. All traces include cache coherency traffic between the L1 cache (compute chiplets), L2 cache (memory chiplets), and main memory (IO chiplets). We use

an adjusted heterogeneous chiplet placement with compute chiplets in the center, memory chiplets on the left and right, and IO chiplets on the top and bottom (see Fig. 9). Each data packet is split into multiple flits, with the number of flits inversely proportional to the topology's link bandwidth. Control packets are modeled as single flits.
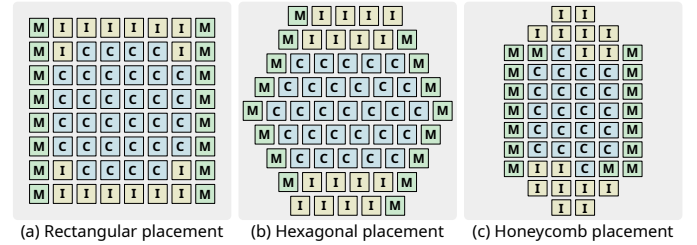


(a) Rectangular placement  (b) Hexagonal placement  (c) Honeycomb placement

Fig. 9: **(§V-E) Placement of compute chiplets (C), memory chiplets (M), and IO chiplets (I) used with traces.**

Fig. 10 shows the throughput and latency of *FoldedHexaTorus* and the baseline topologies on an organic substrate for the five regions of the two traces. We observe that *FoldedHexaTorus* achieves very low, and in some cases the lowest, packet latency while maintaining reasonable throughput.

### VI. MANUFACTURING CONSIDERATIONS

While most ICI topologies assume a rectangular placement of chiplets (see Fig. 6a), we propose using a hexagonal chiplet placement (see Fig. 6b), originally introduced for the *HexaMesh* topology [11]. When applied to systems based on silicon interposers, this hexagonal arrangement results in a mismatch between the rectangular shape of the interposer and the hexagonal layout of the chiplets, leading to a non-negligible interposer area overhead. Although manufacturing hexagonal interposers is technically feasible using plasma dicing [42] or stealth dicing [43], these methods are less widely adopted than conventional blade dicing, which only supports rectangular shapes. In contrast, organic substrates are diced using mechanical routing or laser cutting, both
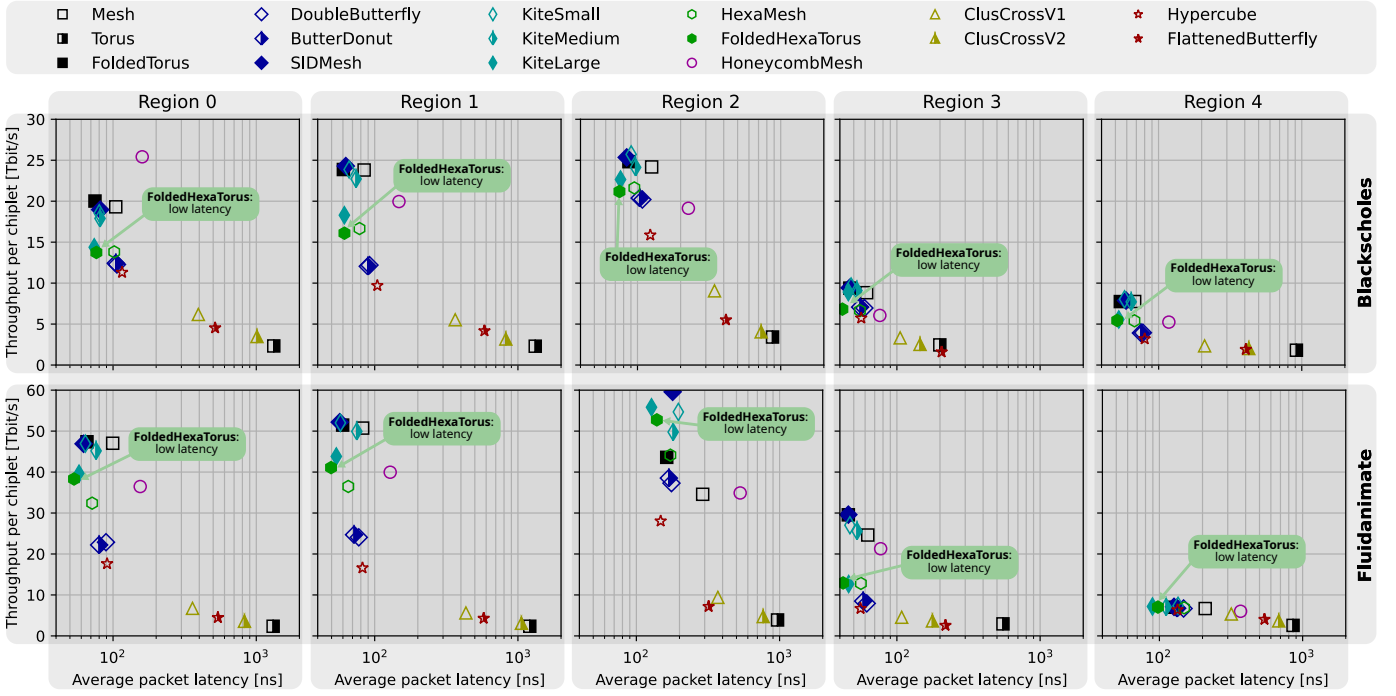
Fig. 10: (**§V-E**) **Evaluation on traffic traces**. Points with saturated colors represent Pareto-optimal topologies.

of which support arbitrary shapes. For glass substrates, mechanical scribbling and breaking or laser cutting are used, with the latter also supporting arbitrary geometries. Moreover, since hexagons tessellate the plane without gaps, there is no substrate material waste when producing hexagonal substrates. In conclusion, while the hexagonal chiplet placement may pose challenges for silicon interposer-based systems, it presents no disadvantages for systems built on organic or glass substrates.

## VII. RELATED WORK

Chiplet-based systems using organic substrates are an established technology. Since AMD [26] demonstrated the cost-efficiency of such systems with their 9-chiplet EPYC processors, the number of chiplets per system has grown significantly, reaching 25 in Tesla's Dojo architecture [28].

Glass substrates, on the other hand, are a more recent technology currently under active development. The new opportunities enabled by glass substrates and their associated challenges are well summarized in the work by Usman et al. [6]. While Vanna-Iampikul et al. [5] highlight the advantages of glass interposers in terms of area efficiency, wire length, signal integrity, and thermal stability, Kim [21] addresses one of their key challenges: power/ground noise.

Regardless of the packaging technology, chiplet-based systems require a high-throughput ICI to provide sufficient communication bandwidth between chiplets. Most prior work has focused on active silicon interposers. Jerger et al. [12] proposed leveraging the interposer's metal layers to implement an ICI that handles most of the core-to-memory traffic, using the *DoubleButterfly* topology. Kannan et al. [7] later introduced the *ButterDonut* topology, disintegrating the compute chiplet

into multiple smaller ones. Further developments include the *ClusCross* [13], *Kite* [14], and *SID-Mesh* [15] topologies.

One of the few works focusing on passive silicon interposers and organic substrates, rather than active interposers, is *HexaMesh* [11], which arranges chiplets in a hexagonal layout and connects each chiplet to its six neighbors. With *FoldedHexaTorus*, we propose— to the best of our knowledge—the first ICI topology optimized for organic and glass substrates.

## VIII. CONCLUSION

Based on our analysis of how network diameter, link-range, and network radix affect an ICI's throughput, latency, area, and power, we define three design principles for ICI topologies on organic and glass substrates: 1) minimize network diameter, 2) use a link-range of one, and 3) minimize network radix. Guided by these principles, we propose the novel *FoldedHexaTorus* topology, which has a link-range of one, a network radix of six, and a network diameter below $\sqrt{N}$, where $N$ is the number of chiplets. We evaluate *FoldedHexaTorus* against a broad set of baseline topologies. Across system sizes from 16 to 256 chiplets and for both organic and glass substrates, it consistently delivers high throughput and near-optimal latency, while incurring only minor area and power overheads.

## References

[1] T. Li, J. Hou, J. Yan, R. Liu, H. Yang, and Z. Sun, "Chiplet heterogeneous integration technology—status and challenges," *Electronics*, 2020.

[2] B. Dehlaghi and A. C. Carusone, "A 0.3 pj/bit 20 gb/s/wire parallel interface for die-to-die communication," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 11, pp. 2690–2701, 2016.

[3] H. Braunisch, A. Aleksov, S. Lotz, and J. Swan, "High-speed performance of silicon bridge die-to-die interconnects," in *2011 IEEE 20th Conference on Electrical Performance of Electronic Packaging and Systems*. IEEE, 2011, pp. 95–98.

[4] Y. Nishi, J. W. Poulton, W. J. Turner, X. Chen, S. Song, B. Zimmer, S. G. Tell, N. Nedovic, J. M. Wilson, W. J. Dally *et al.*, "A 0.297-pj/bit 50.4-gb/s/wire inverter-based short-reach simultaneous bi-directional transceiver for die-to-die interface in 5-nm cmos," *IEEE Journal of Solid-State Circuits*, vol. 58, no. 4, pp. 1062–1073, 2023.

[5] P. Vanna-Iampikul *et al.*, "Glass interposer integration of logic and memory chiplets: Ppa and power/signal integrity benefits," in *2023 60th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2023.

[6] A. Usman *et al.*, "Interposer technologies for high-performance applications," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2017.

[7] A. Kannan, N. E. Jerger, and G. H. Loh, "Enabling interposer-based disintegration of multi-core processors," in *Proceedings of the 48th international symposium on Microarchitecture*, 2015.

[8] R. Mahajan, R. Sankman, N. Patel, D.-W. Kim, K. Aygun, Z. Qian, Y. Mekonnen, I. Salama, S. Sharan, D. Iyengar *et al.*, "Embedded multi-die interconnect bridge (emib)–a high density, high bandwidth packaging interconnect," in *2016 IEEE 66th Electronic Components and Technology Conference (ECTC)*. IEEE, 2016, pp. 557–565.

[9] K. Sikka, R. Bonam, Y. Liu, P. Andry, D. Parekh, A. Jain, M. Bergendahl, R. Divakaruni, M. Cournoyer, P. Gagnon *et al.*, "Direct bonded heterogeneous integration (dbhi) si bridge," in *2021 IEEE 71st Electronic Components and Technology Conference (ECTC)*, 2021, pp. 136–147.

[10] Y. Han *et al.*, "The big chip: Challenge, model and architecture," *Fundamental Research*, 2023.

[11] P. Iff, M. Besta, M. Cavalcante, T. Fischer, L. Benini, and T. Hoefler, "Hexamesh: Scaling to hundreds of chiplets with an optimized chiplet arrangement," in *2023 60th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2023.

[12] N. E. Jerger, A. Kannan, Z. Li, and G. H. Loh, "Noc architectures for silicon interposer systems: Why pay for more wires when you can get them (from your interposer) for free?" in *2014 47th Annual IEEE/ACM International Symposium on Microarchitecture*. IEEE, 2014.

[13] H. Shabani and X. Guo, "Cluscross: a new topology for silicon interposer-based network-on-chip," in *Proceedings of the 13th IEEE/ACM International Symposium on Networks-on-Chip*, 2019.

[14] S. Bharadwaj, J. Yin, B. Beckmann, and T. Krishna, "Kite: A family of heterogeneous interposer topologies enabled via accurate interconnect modeling," in *2020 57th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2020.

[15] B. Sharifpour, M. Sharifpour, and M. Reshadi, "Sid-mesh: Diagonal mesh topology for silicon interposer in 2.5 d noc with introducing a new routing algorithm," in *2021 ACM/IEEE International Workshop on System Level Interconnect Prediction (SLIP)*. IEEE, 2021.

[16] M. Miller and J. Sirán, "Moore graphs and beyond: A survey of the degree/diameter problem," *The electronic journal of combinatorics*, 2012.

[17] G. Haley, "The race to glass substrates," Semiconductor Engineering, 2024, accessed: 2024-08-07. [Online]. Available: https://semiengineering.com/the-race-to-glass-substrates/

[18] E. Alon, M. Hempel, K. Poulton, S. Ardalan, and B. Vinnakota, "Bunch of Wires (BoW) PHY Specification," https://opencomputeproject.github.io/ODSA-BoW/bow_specification.html.

[19] P. Vivet, E. Guthmuller, Y. Thonnart, G. Pillonnet, C. Fuguet, I. Miro-Panades, G. Moritz, J. Durupt, C. Bernard, D. Varreau *et al.*, "Intact: A 96-core processor with six chiplets 3d-stacked on an active interposer with distributed interconnects and integrated power management," *IEEE Journal of Solid-State Circuits*, 2020.

[20] M. N. Sadiku and L. C. Agba, "A simple introduction to the transmission-line modeling," *IEEE Transactions on Circuits and systems*, 1990.

[21] Y. Kim, "Electrical performance analysis of high-speed interconnection and power delivery network (pdn) in low-loss glass substrate-based interposers," *Micromachines*, 2023.

[22] M. Besta and T. Hoefler, "Slim fly: A cost effective low-diameter network topology," in *SC'14: proceedings of the international conference for high performance computing, networking, storage and analysis*. IEEE, 2014.

[23] M. Besta, S. M. Hassan, S. Yalamanchili, R. Ausavarungnirun, O. Mutlu, and T. Hoefler, "Slim noc: A low-diameter on-chip network topology for high energy efficiency and scalability," *ACM SIGPLAN Notices*, 2018.

[24] P. Iff, M. Besta, M. Cavalcante, T. Fischer, L. Benini, and T. Hoefler, "Sparse hamming graph: A customizable network-on-chip topology," in *2023 60th ACM/IEEE Design Automation Conference (DAC)*, 2023.

[25] J. Kim, W. J. Dally, S. Scott, and D. Abts, "Technology-driven, highly-scalable dragonfly topology," *ACM SIGARCH Computer Architecture News*.

[26] S. Naffziger, N. Beck, T. Burd, K. Lepak, G. H. Loh, M. Subramony, and S. White, "Pioneering chiplet technology and design for the amd epyc™ and ryzen™ processor families: Industrial product," in *2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2021.

[27] The UCIe Consortium, "Universal Chiplet Interconnect Express (UCIe) Specification," https://www.uciexpress.org/specification.

[28] E. Talpes, D. D. Sarma, D. Williams, S. Arora, T. Kunjan, B. Floering, A. Jalote, C. Hsiong, C. Poorna, V. Samant *et al.*, "The microarchitecture of dojo, tesla's exa-scale computer," *IEEE Micro*, 2023.

[29] P.-H. Pham, P. Mau, and C. Kim, "A 64-pe folded-torus intra-chip communication fabric for guaranteed throughput in network-on-chip based applications," in *2009 IEEE Custom Integrated Circuits Conference*. IEEE, 2009.

[30] A. Sahba and J. J. Prevost, "Hypercube based clusters in cloud computing," in *2016 World Automation Congress (WAC)*. IEEE, 2016.

[31] J. Kim, J. Balfour, and W. Dally, "Flattened butterfly topology for on-chip networks," in *40th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO 2007)*. IEEE, 2007.

[32] I. Stojmenovic, "Honeycomb networks: Topological properties and communication algorithms," *IEEE Transactions on parallel and distributed systems*, 1997.

[33] N. Jiang, D. U. Becker, G. Michelogiannakis, J. Balfour, B. Towles, D. E. Shaw, J. Kim, and W. J. Dally, "A detailed and flexible cycle-accurate network-on-chip simulator," in *2013 IEEE international symposium on performance analysis of systems and software (ISPASS)*. IEEE, 2013.

[34] C. J. Glass and L. M. Ni, "The turn model for adaptive routing," *ACM SIGARCH Computer Architecture News*, vol. 20, no. 2, 1992.

[35] L. Levitin, M. Karpovsky, and M. Mustafa, "Deadlock prevention by turn prohibition in interconnection networks," in *2009 IEEE international symposium on parallel & distributed processing*. IEEE, 2009.

[36] T. Caldwell, "On finding minimum routes in a network with turn penalties," *Communications of the ACM*, vol. 4, no. 2, pp. 107–108, 1961.

[37] J. Hestness, B. Grot, and S. W. Keckler, "Netrace: dependency-driven trace-based network-on-chip simulation," in *Proceedings of the Third International Workshop on Network on Chip Architectures*, 2010, pp. 31–36.

[38] J. Hestness and S. W. Keckler, "Netrace: Dependency-tracking traces for efficient network-on-chip experimentation," *The University of Texas at Austin, Dept. of Computer Science, Tech. Rep*, 2011.

[39] P. Iff, B. Bruggmann, M. Besta, L. Benini, and T. Hoefler, "Rapidchiplet: A toolchain for rapid design space exploration of chiplet architectures," *arXiv preprint arXiv:2311.06081*, 2023.

[40] J. Hestness, B. Grot, and S. W. Keckler, "Netraces v1.0 (A collection of network traces with dependency information)." https://www.cs.utexas.edu/ netrace/.

[41] C. Bienia, S. Kumar, J. P. Singh, and K. Li, "The parsec benchmark suite: Characterization and architectural implications," in *Proceedings of the 17th international conference on Parallel architectures and compilation techniques*, 2008, pp. 72–81.

[42] N. Matsubara, R. Windemuth, H. Mitsuru, and H. Atsushi, "Plasma dicing technology," in *2012 4th Electronic System-Integration Technology Conference*. IEEE, 2012, pp. 1–5.

[43] M. Kumagai, N. Uchiyama, E. Ohmura, R. Sugiura, K. Atsumi, and K. Fukumitsu, "Advanced dicing technology for semiconductor wafer—stealth dicing," *IEEE Transactions on Semiconductor Manufacturing*, vol. 20, no. 3, pp. 259–265, 2007.